上海市研究生教育专项经费资助项目

凸分析与非光滑分析

胡毓达　孟志青　著

上 海 科 学 技 术 出 版 社

图书在版编目(CIP)数据

凸分析与非光滑分析/胡毓达,孟志青著.—上海:上海科学技术出版社,2000.9
上海研究生教育用书
ISBN 7-5323-5525-X

Ⅰ.凸… Ⅱ.①胡…②孟… Ⅲ.凸分析-研究生-教材 Ⅳ.O174.13

中国版本图书馆 CIP 数据核字(2000)第 41258 号

责任编辑 赵序明

上海科学技术出版社出版发行
(上海瑞金二路 450 号 邮政编码 200020)
上海新华印刷厂印刷 新华书店上海发行所经销
2000 年 9 月第 1 版 2000 年 9 月第 1 次印刷
开本 850×1156 1/32 印张 12 插页 4 字数 312 000
印数 1-1 500 定价:25.00 元

内容提要

本书论述凸分析与非光滑分析的基本理论，特别是较系统地介绍了有关集值映射的广义微分的研究进展．全书共 8 章：第 1 章至第 3 章是凸分析部分，包括凸集、凸函数和凸映射以及凸函数的次微分；第 4 章概述非光滑的 Lipschitz 函数的广义次微分；第 5 章则利用切锥引进并讨论函数的切导数和切微分；第 6 章至第 8 章研究集值映射的广义微分，其中第 6 章考虑锥凸集值映射的锥次微分；第 7 章进而讨论一般集值映射的锥弱次微分；第 8 章则从另一角度研究集值映射的切导数和切微分．

本书适用于作大学硕士和博士研究生的教材或主要教学参考书，也可供大学本科高年级学生选读．同时，它也是从事现代分析、数学规划、向量极值、最优控制和数理经济等领域研究的学者的重要参考书．

前　言

凸分析和非光滑分析是20世纪60年代至80年代相继发展形成的现代数学分支．作为描述和解决非光滑问题的有力工具，它们在非线性最优化、多目标决策、最优控制、对策论、变分学、逼近理论以及数理经济等领域有着广泛的应用．事实上，凸分析主要研究的凸函数（凸泛函）在通常意义下是非光滑的，因此从这一意义上说，凸分析也是非光滑分析的组成部分和重要基础，而非光滑分析则是凸分析研究的延伸和发展．

本书是我自1986年至1998年间在上海交通大学先后为应用数学专业的硕士和博士研究生讲授凸分析、非光滑分析和现代分析课的基础上，经过选择整理写成的．全书脱稿后，于2000年2月至6月在温州大学对该校和温州师范学院的数学教师以及访问学者试讲了部分内容，并稍作修改。

书中的第1章至第3章是相对成熟的凸分析部分，第2章中有关映射的拟凸性的内容，主要是我和学生胡一凡博士以及访问学者凌晨副教授合作的工作．第4章关于Lipschitz函数的广义次微分和第5章函数的切微分，部分取材于F. H. Clarke的著作[11]．第5章和第8章的部分内容，参考了J. P. Aubin和I. Ekeland的书[16]，以及B. S. Morduckhovich的论文[28]．第6章至第8章是集值映射的广义微分部分，其基本概念先是由我和学生徐永明博士开始引进，而后则是和孟志青博士共同进行了较系统研究的结果．本书的第1章至第3章由我执笔撰写．1994年至1995年间，湘潭大学数学系孟志青副教授作为访问学者来我处进修，他除了与我合作进行有关集值映射的广义微分的研究之外，还曾在我讲授的非光滑分析课中分讲了部分内容．为此，本书的第4章至第8章先

由他整理出初稿.经我们多次讨论调整,最后由我作修改、补充和统一定稿.

阅读本书一般需要有泛函分析的基本知识.特别是从第5章以后,读者最好能熟悉非线性泛函的某些近代理论,同时对其中一些比较简洁的推证要进行仔细地研读。

在本书的写作过程中,陆晋奎教授、王晓敏副教授和杨雷博士曾协助我补充和查对了有关资料,丁鸿生讲师和于丽英博士则为打印本书的手稿付出了辛勤的劳动.在此,一并致以衷心感谢!

限于作者水平,书中难免存在不妥和谬误之处,恳望专家和读者不吝批评指正.

胡毓达

2000年1月1日于上海交通大学

2000年7月7日补写于温州大学

符号说明

符号	说明
$\mathscr{V}$，$\mathscr{W}$	线性空间
$\mathscr{L}$	线性子空间
$\mathscr{X}$，$\mathscr{Y}$，$\mathscr{Z}$	线性拓扑空间
$\mathscr{X}^*$，$\mathscr{Y}^*$，$\mathscr{Z}^*$	线性拓扑空间 $\mathscr{X}$，$\mathscr{Y}$，$\mathscr{Z}$ 的对偶空间
$\mathscr{H}$	Hilbert 空间
$\mathscr{B}$，$\mathscr{D}$，$\mathscr{B}_1$，$\mathscr{B}_2$	Banach 空间
$\mathscr{B}^*$	Banach 空间 $\mathscr{B}$ 的对偶空间
R^n，R^m	n 维，m 维 Euclid 空间
R	实数域
R_+	非负实数域
S，T，M	集合
$\mathrm{co}S$	集合 S 的凸包
$\mathrm{aff}S$	集合 S 的仿射包
$\mathrm{lin}S$	集合 S 的线性包
S^i	集合 S 的代数内部
S^{ri}	集合 S 的相对代数内部
S^a	集合 S 的代数闭包
S^b	集合 S 的代数边界
$\mathrm{int}S$	集合 S 的内部
$\mathrm{ri}S$	集合 S 的相对内部
$\mathrm{cl}S$，$\overline{S}$	集合 S 的闭包
∂S	集合 S 的边界
$\boldsymbol{d}\perp S$	Banach 空间中向量 $\boldsymbol{d}$ 垂直于集合 S
$S^{\perp}$	Banach 空间中集合 S 的垂直集

符号	说明
H	超平面
H_+^C, H_-^C	由超平面 H 界定的闭半空间
H_+^0, H_-^0	由超平面 H 界定的开半空间
$U(\boldsymbol{x})$	线性拓扑空间中点 $\boldsymbol{x}$ 的邻域
U, $U(\boldsymbol{O})$	线性拓扑空间中原点的邻域
$U(\boldsymbol{x}, \delta)$	Banach 空间中点 $\boldsymbol{x}$ 的 δ-邻域
$B_\delta(\boldsymbol{x})$	Banach 空间中以点 $\boldsymbol{x}$ 为中心，δ 为半径的超球
B	Banach 空间中的开单位球(或闭单位球)
$(\boldsymbol{x}^1, \boldsymbol{x}^2)$	线性空间中点 $\boldsymbol{x}^1$ 与 $\boldsymbol{x}^2$ 之间的开线段
$[\boldsymbol{x}^1, \boldsymbol{x}^2]$	线性空间中点 $\boldsymbol{x}^1$ 与 $\boldsymbol{x}^2$ 之间的闭线段
$(\boldsymbol{x}^1, \boldsymbol{x}^2]$, $[\boldsymbol{x}^1, \boldsymbol{x}^2)$	线性空间中点 $\boldsymbol{x}^1$ 与 $\boldsymbol{x}^2$ 之间的半开半闭线段
$\|\cdot\|_{\mathscr{B}}$, $\|\cdot\|$	Banach 空间上的范数
$\|\cdot\|_p$	Euclid 空间上的 p-范数
$\|M\|_{m\times n}$	$m\times n$ 阶矩阵 M 的范数
K	线性空间中的锥
K^*	锥 K 的对偶锥
K^{**}	锥 K 的双对偶锥
K^+	锥 K 的严格对偶锥
K^-	锥 K 的负对偶锥
R_+^n	n 维 Euclid 空间中的非负锥
S^-	集合 S 的负对偶集(极化集)
cone$S$	集合 S 的锥包
f, g	线性空间上的实值函数(实泛函)
$\boldsymbol{f}$	线性空间上的实值向量函数
$\langle \boldsymbol{x}^*, \boldsymbol{x}\rangle$	线性空间上线性泛函 $\boldsymbol{x}^*$ 在点 $\boldsymbol{x}$ 处的值
φ	线性拓扑空间上的映射
φ^*	线性映射 φ 的对偶算子
ψ	线性拓扑(Banach)空间上的集值映射

L	线性空间到线性空间的线性映射
$L(\mathscr{X}, \mathscr{Y})$	线性拓扑空间 $\mathscr{X}$ 到 $\mathscr{Y}$ 的连续线性映射(算子)集合
$\mathrm{dom} f$	广义实值函数 f 的有效域
$\mathrm{epi} f$	实值函数 f 的上图象
$\mathrm{hyp} f$	实值函数 f 的下图象
$\mathrm{co}(f)$	实值函数 f 的凸包(函数)
$\mathrm{cl}(f)$	实值函数 f 的闭包(函数)
$\liminf\limits_{x \to x^0} f(\boldsymbol{x})$	实值函数 f 在点 $\boldsymbol{x}^0$ 处的下极限
$\limsup\limits_{x \to x^0} f(\boldsymbol{x})$	实值函数 f 在点 $\boldsymbol{x}^0$ 处的上极限
$\mathrm{dom}\psi$	集值映射 ψ 的有效域
$\mathrm{graph}\psi$	集值映射 ψ 的图象
$K\text{-epi}\psi$	集值映射 ψ 的 K-上图象
$\mathrm{Ker}\psi$	集值映射 ψ 的核
$\mathrm{Im}\psi$	集值映射 ψ 的象
$H_S(f, c)$	集合 S 上函数 f 关于 c 的水平集
$H_S(\varphi, c)_K$	集合 S 上映射 φ 关于 c 的 K-水平集
$K\text{-bou}\{\boldsymbol{y}^1, \cdots, \boldsymbol{y}^m\}$	线性拓扑空间中向量组 $\{\boldsymbol{y}^1, \cdots, \boldsymbol{y}^m\}$ 的 K-有界集
f^*	凸函数 f 的共轭函数
f^{**}	凸函数 f 的二次共轭函数
g^*	凹函数 g 的共轭函数
d_S, $d_S(\boldsymbol{x})$	Banach 空间中集合 S 上的距离函数
μ_S, $\mu_S(\boldsymbol{x})$	线性拓扑空间中凸集 S 上的 Minkowski 函数(泛函)
μ_a, $\mu_a(\boldsymbol{y}, \boldsymbol{\omega})$	$\mathscr{Y} \times \mathrm{int}K$ 上关于 $a \in \mathscr{Y}$ 的 Minkowski 泛函
δ_S, $\delta_S(\boldsymbol{x})$	线性空间中凸集 S 上的指示函数
σ_S, $\sigma_S(\boldsymbol{x})$	线性空间的对偶空间中集合 S 的支撑函数
l, $l(\boldsymbol{x})$	线性空间上的仿射函数

$T_S(\boldsymbol{x}^0)$　线性拓扑空间中凸集 S 在点 $\boldsymbol{x}^0$ 处的切锥

$N_S(\boldsymbol{x}^0)$　线性拓扑空间中凸集 S 在点 $\boldsymbol{x}^0$ 处的法锥

$T_S^0(\boldsymbol{x}^0)$　Banach 空间中集合 S 在点 $\boldsymbol{x}^0$ 处的 G-切锥（广义切锥）

$N_S^0(\boldsymbol{x}^0)$　Banach 空间中集合 S 在点 $\boldsymbol{x}^0$ 处的 G-法锥（广义法锥）

$Q_S(\boldsymbol{x}^0)$　Banach 空间中集合 S 在点 $\boldsymbol{x}^0$ 处的相依锥

$T_S^H(\boldsymbol{x}^0)$　Banach 空间中集合 S 在点 $\boldsymbol{x}^0$ 处的 H-切锥（超切锥）

$T_S^D(\boldsymbol{x}^0)$　Banach 空间中集合 S 在点 $\boldsymbol{x}^0$ 处的 D-切锥（相依切锥）

$T_S^C(\boldsymbol{x}^0)$　Banach 空间中集合 S 在点 $\boldsymbol{x}^0$ 处的 C-切锥（Clarke 切锥）

$N_S^C(\boldsymbol{x}^0)$　Banach 空间中集合 S 在点 $\boldsymbol{x}^0$ 处的 C-法锥（Clarke 法锥）

$G_{K\text{-epi}\psi}(\boldsymbol{x}^0, \boldsymbol{y}^0)$　集合 K-epi ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的 G-切锥

$N_S^M(\boldsymbol{x}^0)$　集合 S 在点 $\boldsymbol{x}^0$ 处的 M-法锥

$N_S^F(\boldsymbol{x}^0, \varepsilon)$　集合 S 在点 $\boldsymbol{x}^0$ 处的 F-法锥(Frechet-ε 法锥)

$N_{K\text{-epi}\psi}(x^0, y^0)$　集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的法锥

$\mathscr{E}(Y, K)$　线性拓扑空间中集合 Y 的 K-有效点集

$\mathscr{E}_w(Y, K)$　线性拓扑空间中集合 Y 的 K-弱有效点集

$\nabla f(\boldsymbol{x}^0)$　Euclid 空间上实值函数 f 在点 $\boldsymbol{x}^0$ 处的梯度

$\nabla^2 f(\boldsymbol{x}^0)$　Euclid 空间上实值函数 f 在点 $\boldsymbol{x}^0$ 处的 Hesse 矩阵

$J\boldsymbol{f}(\boldsymbol{x}^0)$　Euclid 空间上实值向量函数 $\boldsymbol{f}$ 在点 $\boldsymbol{x}^0$ 处的 Jacobi 矩阵

$f'(\boldsymbol{x}^0)$　Banach 空间上实值函数 f 在点 $\boldsymbol{x}^0$ 处的 Fréchet 导数

φ'_{x^0}	线性拓扑空间上映射 φ 在点 $\boldsymbol{x}^0$ 处的 Gâteaux 导数
$D_S\varphi(\boldsymbol{x}^0)$	Banach 空间上映射 φ 在点 $\boldsymbol{x}^0$ 处的严格导数
$f'_+(\boldsymbol{x}^0;\boldsymbol{d})$	实值函数 f 在点 $\boldsymbol{x}^0$ 处沿方向 $\boldsymbol{d}$ 的右方向导数
$f'_-(\boldsymbol{x}^0;\boldsymbol{d})$	实值函数 f 在点 $\boldsymbol{x}^0$ 处沿方向 $\boldsymbol{d}$ 的左方向导数
$f'(\boldsymbol{x}^0;\boldsymbol{d})$	实值函数 f 在点 $\boldsymbol{x}^0$ 处沿方向 $\boldsymbol{d}$ 的方向导数
$f^0(\boldsymbol{x}^0;\boldsymbol{d})$	Lipschitz 函数 f 在点 $\boldsymbol{x}^0$ 处沿方向 $\boldsymbol{d}$ 的 G-方向导数(广义方向导数)
$f^D(\boldsymbol{x}^0;\boldsymbol{d})$	正常函数 f 在点 $\boldsymbol{x}^0$ 处沿方向 $\boldsymbol{d}$ 的 D 切导数
$f^C(\boldsymbol{x}^0;\boldsymbol{d})$	正常函数 f 在点 $\boldsymbol{x}^0$ 处沿方向 $\boldsymbol{d}$ 的 C-切导数(上图象导数)
$f^H(\boldsymbol{x}^0;\boldsymbol{d})$	实值函数 f 在点 $\boldsymbol{x}^0$ 处沿方向 $\boldsymbol{d}$ 的 H 切导数
$\partial f(\boldsymbol{x}^0)$	凸函数 f 在点 $\boldsymbol{x}^0$ 处的次微分
$\partial^0 f(\boldsymbol{x}^0)$	函数 f 在点 $\boldsymbol{x}^0$ 处的 G-次微分
$\partial_1^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2)$	Lipschitz 函数 f 在点 $\boldsymbol{x}_1$ 处的 G-偏次微分
$\partial_2^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2)$	Lipschitz 函数 f 在点 $\boldsymbol{x}_2$ 处的 G-偏次微分
$\partial^0 \boldsymbol{f}(\boldsymbol{x}^0)$	Lipschitz 向量函数 $\boldsymbol{f}$ 在点 $\boldsymbol{x}^0$ 处的 G-次微分
$\partial^C f(\boldsymbol{x}^0)$	正常函数 f 在点 $\boldsymbol{x}^0$ 处的 C-切微分
$\partial^M f(\boldsymbol{x}^0)$	广义实值函数 f 在点 $\boldsymbol{x}^0$ 处的 M 次微分
$\partial^\infty f(\boldsymbol{x}^0)$	广义实值函数 f 在点 $\boldsymbol{x}^0$ 处的 M-奇异次微分
$\psi'(\boldsymbol{x}^0, \boldsymbol{y}^0;\boldsymbol{d})$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处沿方向 $\boldsymbol{d}$ 的 K-方向导数集合
$\psi'_w(\boldsymbol{x}^0, \boldsymbol{y}^0;\boldsymbol{d})$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处沿方向 $\boldsymbol{d}$ 的 K-弱方向导数集合
$\psi^0_w(\boldsymbol{x}^0, \boldsymbol{y}^0;\boldsymbol{d})$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处沿方向 $\boldsymbol{d}$ 的 GK-弱方向导数集合

$\psi_w^C(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d})$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处沿方向 $\boldsymbol{d}$ 的 CK-弱方向导数集合
$\psi^T(\boldsymbol{x}^0, \boldsymbol{y}^0)$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的切导数
$\psi^D(\boldsymbol{x}^0, \boldsymbol{y}^0)$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的 D-切导数
$\psi^C(\boldsymbol{x}^0, \boldsymbol{y}^0)$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的 C-切导数
$\partial\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)_{\boldsymbol{p}}$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处关于 $\boldsymbol{p}$ 的 K-次微分
$\partial_w\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)_{\boldsymbol{p}}$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处关于 $\boldsymbol{p}$ 的 K-弱次微分
$\partial\psi(\boldsymbol{x}^0)_{\boldsymbol{p}}$	集值映射 ψ 在点 $\boldsymbol{x}^0$ 处关于 $\boldsymbol{p}$ 的 K-次微分
$\partial_w\psi(\boldsymbol{x}^0)_{\boldsymbol{p}}$	集值映射 ψ 在点 $\boldsymbol{x}^0$ 处关于 $\boldsymbol{p}$ 的 K-弱次微分
$\partial_w^0\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)_{\boldsymbol{p}}$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处关于 $\boldsymbol{p}$ 的 GK-弱次微分
$\partial_w^C\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)_{\boldsymbol{p}}$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处关于 $\boldsymbol{p}$ 的 CK-弱次微分
$\partial^T\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的切微分
$\partial^C\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的 C-切微分
$\partial\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)_L$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的算子 K-次微分
$\partial_w\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)_L$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的算子 K-弱次微分
$\partial_w^0\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)_L$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的算子 GK-弱次微分
$\partial_w^0\psi(\boldsymbol{x}^0)_L$	集值映射 ψ 在点 $\boldsymbol{x}^0$ 处的算子 GK-弱次微分
$\partial_w^C\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)_L$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的算子 CK-弱次微分
$\partial_w^C\psi(\boldsymbol{x}^0)_L$	集值映射 ψ 在点 $\boldsymbol{x}^0$ 处的算子 CK-弱次微分
$\partial^M\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的 M-次微分
$\partial^{MC}\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的 MC-次微分
$\partial^{MF}\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)$	集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的 MF-次微分

目　　录

第1章 凸　　集

凸集是凸分析的研究对象和基本内容.它的有关概念和理论，则是凸分析和非光滑分析的基础和重要研究工具.

这一章将论述凸集及有关的概念和性质.特别是，介绍有着广泛应用的在不同空间情况的几个凸集分离定理.同时还讨论作为重要特殊凸集的凸锥及其对偶锥.

§1.1　凸集及有关性质

本节阐述具有所谓凸性的集合及有关概念，并且介绍它们的基本性质.

设 $\mathscr{X}$ 是实线性空间，它不止含有一个点.

定义 1.1.1　设集合 $S\subset\mathscr{X}$.若对任意的 $\lambda\in(0,1)$ 有

$$\lambda \boldsymbol{x}^1+(1-\lambda)\boldsymbol{x}^2\in S\quad\forall \boldsymbol{x}^1,\boldsymbol{x}^2\in S,$$

则称 S 是($\mathscr{X}$ 中的)凸集或集合 S 是凸的.

我们约定：空集是凸集.

图 1.1.1 的(a)和(b)分别是凸集和不是凸集的图示.

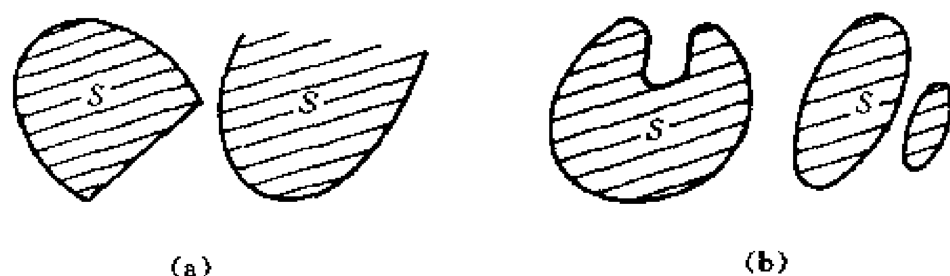


图　1.1.1

例 1.1.1 以下集合是相应空间中的凸集.

Euclid 空间 R^n 中的超平面:

$$H = \{x \in R^n \mid a^T x = \beta, a \in R^n, \beta \in R\}.$$

Euclid 空间 R^n 中的开半空间:

$$H_+^0 = \{x \in R^n \mid a^T x > \beta, a \in R^n, \beta \in R\}.$$

Banach 空间 $\mathscr{B}$ 中以 x^0 为中心,δ 为半径的超球:

$$B_\delta(x^0) = \{x \in \mathscr{B} \big| \| x - x^0 \| < \delta, x^0 \in \mathscr{B}, \delta > 0\}.$$

特别地,$\mathscr{B}$ 中的开单位球:

$$B = \{x \in \mathscr{B} \big| \| x \| < 1\},$$

其中 $\| x \|$ 是 x 的范数.

线性空间 $\mathscr{V}$ 中的闭线段:

$$[x^1, x^2] = \{x \in \mathscr{V} \mid x = \lambda x^1 + (1-\lambda)x^2,$$

$$x^1, x^2 \in \mathscr{V}, \lambda \in [0, 1]\}.$$

线性空间 $\mathscr{V}$ 中的正系数多项式集合:

$$P = \{p \in \mathscr{V} \mid p \text{ 是 } R \text{ 上正系数单变量多项式}\}.$$

先叙述关于凸集的一些简单性质.

定理 1.1.1 设集合 $S \subset \mathscr{V}$. S 是凸集当且仅当

$$(\alpha_1 + \alpha_2)S = \alpha_1 S + \alpha_2 S \quad \forall \alpha_1, \alpha_2 \in R_+. \tag{1.1.1}$$

证明 必要性. 当 $\alpha_1 = \alpha_2 = 0$ 时,结论显然成立. 下设 $\alpha_1 + \alpha_2 > 0$.

先证等式(1.1.1)的右端包含左端. 为此,设 $z \in (\alpha_1 + \alpha_2)S$,则存在 $x \in S$ 使 $z = (\alpha_1 + \alpha_2)x = \alpha_1 x + \alpha_2 x$. 因为 $\alpha_1 x \in \alpha_1 S$, $\alpha_2 x \in \alpha_2 S$, 故 $z \in \alpha_1 S + \alpha_2 S$. 为证(1.1.1)的左端包含右端,设 $z \in$

$\alpha_1 S + \alpha_2 S$，则存在 x^1，$x^2 \in S$ 使

$$z = \alpha_1 x^1 + \alpha_2 x^2 = (\alpha_1 + \alpha_2)\left(\frac{\alpha_1}{\alpha_1 + \alpha_2}x^1 + \frac{\alpha_2}{\alpha_1 + \alpha_2}x^2\right).$$

因为 S 是凸的，由定义 1.1.1 有 $\frac{\alpha_1}{\alpha_1 + \alpha_2}x^1 + \frac{\alpha_2}{\alpha_1 + \alpha_2}x^2 \in S$，于是得 $z \in (\alpha_1 + \alpha_2)S$.

充分性. 已知(1.1.1)成立. 对于任意的 x^1，$x^2 \in S$ 和任意的 $\lambda \in (0, 1)$，任取 $\alpha_1 > 0$，令 $\alpha_2 = \frac{1-\lambda}{\lambda}\alpha_1$，则有 $\lambda = \frac{\alpha_1}{\alpha_1 + \alpha_2}$，$1 - \lambda = \frac{\alpha_2}{\alpha_1 + \alpha_2}$. 由(1.1.1)可知 $\alpha_1 x^1 + \alpha_2 x^2 \in \alpha_1 S + \alpha_2 S = (\alpha_1 + \alpha_2)S$，故存在 $x \in S$ 有 $\alpha_1 x^1 + \alpha_2 x^2 = (\alpha_1 + \alpha_2)x$ 或

$$\frac{\alpha_1}{\alpha_1 + \alpha_2}x^1 + \frac{\alpha_2}{\alpha_1 + \alpha_2}x^2 = x.$$

由此，我们有 $\lambda x^1 + (1-\lambda)x^2 = \frac{\alpha_1}{\alpha_1 + \alpha_2}x^1 + \frac{\alpha_2}{\alpha_1 + \alpha_2}x^2 = x \in S$，于是由定义 1.1.1 得知 S 是凸的. ▯

定理 1.1.2 设 $S_i \subset \mathscr{V}$ ($i \in I$，I 是指标集)是凸集.

(1) 若 $I = \{1, \cdots, m\}$ 是有限指标集，$\alpha_i \in R$ ($i = 1, \cdots, m$)，则 $\sum_{i=1}^{m} a_i S_i$ 是凸集.

(2) 若 I 是任意指标集，则 $\bigcap_{i \in I} S_i$ 是凸集.

证明 由定义 1.1.1 直接可以推得. ▯

定理 1.1.3 设 $\mathscr{W}$ 是线性空间，$L: \mathscr{V} \to \mathscr{W}$ 是线性映射.

(1) 若 $S \subset \mathscr{V}$ 是凸集，则 $L(S) \subset \mathscr{W}$ 是凸集.

(2) 若 $T \subset \mathscr{W}$ 是凸集，则 $L^{-1}(T) \subset \mathscr{V}$ 是凸集.

证明 (1) 任取 y^1，$y^2 \in L(S)$，则存在 x^1，$x^2 \in S$，有 $y^1 = L(x^1)$，$y^2 = L(x^2)$. 对于任意的 $\lambda \in (0, 1)$，因为 S 是凸的，故 $\lambda x^1 + (1-\lambda)x^2 \in S$. 于是，由 L 是线性映射，我们有

$$\lambda y^1 + (1-\lambda)y^2 = \lambda L(x^1) + (1-\lambda)L(x^2)$$

$$= L(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \in L(S).$$

由于 $\boldsymbol{y}^1$ 和 $\boldsymbol{y}^2$ 是任取的，根据定义 1.1.1 得知 $L(S)$是凸的.

（2）由于 $L^{-1}: \mathscr{W} \to \mathscr{V}$ 也是线性映射，与（1）的证明类似得证. ▯

定义 1.1.2 设 $\boldsymbol{x}^i \in \mathscr{V}$, $\lambda_i \geqslant 0$ $(i = 1, \cdots, m)$, $\sum_{i=1}^{m}\lambda_i = 1$, 则称向量 $\boldsymbol{x} = \sum_{i=1}^{m}\lambda_i \boldsymbol{x}^i$ 是$\boldsymbol{x}^1, \cdots, \boldsymbol{x}^m$ 的凸组合.

例 1.1.2 任一圆内的点都是圆周上某两点的凸组合. 任一三角形中的点都是它的三个顶点的凸组合.

定理 1.1.4 设集合 $S \subset \mathscr{V}$ 非空. S 是凸集当且仅当对任何正整数 $m \geqslant 2$ 有 $\boldsymbol{x}^i \in S$, $\lambda_i \geqslant 0$ $(i = 1, \cdots, m)$, $\sum_{i=1}^{m}\lambda_i = 1$, 使得

$$\sum_{i=1}^{m}\lambda_i \boldsymbol{x}^i \in S. \tag{1.1.2}$$

证明 充分性. 在(1.1.2)中取 $m = 2$, 由定义 1.1.1 即知 S 是凸的.

必要性. 用数学归纳法证明：当 $m = 2$ 时，由 S 是凸的，按定义 1.1.1 可知(1.1.2)成立.

设已知 $m = k$ 时(1.1.2)成立. 令

$$\boldsymbol{z} = \sum_{i=1}^{k+1}\lambda_i \boldsymbol{x}^i, \tag{1.1.3}$$

其中 $\lambda_i \geqslant 0$ $(i = 1, \cdots, k+1)$, $\sum_{i=1}^{k+1}\lambda_i = 1$. 若 $\lambda_{k+1} = 1$, 则 $\boldsymbol{z} = \boldsymbol{x}^{k+1} \in S$, 即(1.1.2)对于 $m = k+1$ 成立. 若 $\lambda_{k+1} \neq 1$, 因为(1.1.2)对于 $m = k$ 成立，故有

$$\boldsymbol{z}' = \sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} \boldsymbol{x}^i \in S,$$

$$\frac{\lambda_i}{1-\lambda_{k+1}} \geqslant 0 \ (i=1, \cdots, k), \ \sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} = 1.$$

由此，根据(1.1.3)和 S 是凸的，得到

$$z = \sum_{i=1}^{k} \lambda_i x^i + \lambda_{k+1} x^{k+1} = (1-\lambda_{k+1}) z' + \lambda_{k+1} x^{k+1} \in S,$$

再由(1.1.3)便知(1.1.2)对于 $m = k+1$ 成立. ▯

定理 1.1.4 表明 S 是凸集的充分必要条件是：该集合中任意个点的凸组合仍属于此集合.

以下两个概念是凸集的扩展.

定义 1.1.3 设集合 $S \subset \mathscr{Y}$.

(1) 若对任意的 $\lambda \in R$ 有

$$\lambda x^1 + (1-\lambda) x^2 \in S \quad \forall x^1, x^2 \in S,$$

则称 S 是($\mathscr{Y}$ 中的)仿射集(仿射流形，线性流形)，或集合 S 是仿射的.

(2) 若对任意的 $\lambda_1, \lambda_2 \in R$ 有

$$\lambda_1 x^1 + \lambda_2 x^2 \in S \quad \forall x^1, x^2 \in S,$$

则称 S 是($\mathscr{Y}$ 中的)线性子空间.

我们约定：空集是仿射集.

注 1.1.1 由定义 1.1.1 和定义 1.1.3 可知，仿射集和线性子空间都是凸集. 又集合 S 是 $\mathscr{Y}$ 中的线性子空间，即 S 对 $\mathscr{Y}$ 中的线性运算构成线性空间.

例 1.1.3 Euclid 空间 R^n 中的点、直线和超平面都是 R^n 中的仿射集. R^n 中过原点的直线和过原点的超平面都是 R^n 中的线性子空间.

定理 1.1.5 设集合 $S_i \subset \mathscr{Y}$ ($i \in I$, I 是任意指标集).

(1) 若 S_i ($i \in I$)是仿射集，则 $\bigcap_{i \in I} S_i$ 是仿射集.

(2) 若 S_i ($i \in I$)是线性子空间，则 $\bigcap_{i \in I} S_i$ 是线性子空间.

证明　由定义1.1.3不难直接推得.▯

定义 1.1.4　设点 $x^i \in \mathscr{V}$ $(i=1,\cdots,m)$.

(1) 若 $\lambda_i \in R$ $(i=1,\cdots,m)$, $\sum_{i=1}^{m}\lambda_i=1$, 则称向量 $x=\sum_{i=1}^{m}\lambda_i x^i$ 是 $x^1,\cdots,x^m$ 的仿射组合.

(2) 若 $\lambda_i \in R$ $(i=1,\cdots,m)$, 则称向量 $x=\sum_{i=1}^{m}\lambda_i x^i$ 是 $x^1,\cdots,x^m$ 的线性组合.

定理 1.1.6　设集合 $S\subset\mathscr{V}$ 非空.

(1) S 是仿射集当且仅当对任何正整数 $m\geqslant 2$, 有 $x^i\in S$, $\lambda_i\in R$ $(i=1,\cdots,m)$, $\sum_{i=1}^{m}\lambda_i=1$, 使得 $\sum_{i=1}^{m}\lambda_i x^i\in S$.

(2) S 是线性子空间当且仅当对任何正整数 $m\geqslant 2$ 有 $x^i\in S$, $\lambda_i\in R$ $(i=1,\cdots,m)$, 使得 $\sum_{i=1}^{m}\lambda_i x^i\in S$.

证明　与定理1.1.4的证明类似.▯

现在引进一般集合的凸包概念,并讨论有关的一些性质.

定义 1.1.5　设集合 $S\subset\mathscr{V}$, 则 $\mathscr{V}$ 中所有包含 S 的凸集的交称为 S 的凸包(或由 S 生成的凸集),记作 co$S$.

我们约定:空集的凸包是空集.

注 1.1.2　由定义1.1.5,集合的凸包是包含它的最小凸集(见图1.1.2).

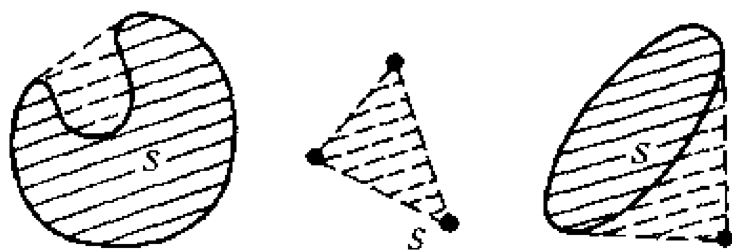


图　1.1.2

定理 1.1.7　设集合 $S\subset\mathscr{V}$.

(1) 若 S 是凸集，则 $\mathrm{co}S = S$.

(2) 若 $\boldsymbol{x} \in \mathrm{co}S$，则 $\mathrm{co}\left(S \cup \{\boldsymbol{x}\}\right) = \mathrm{co}S$.

证明 (1) 由定义 1.1.5 直接可知.

(2) 由 $\boldsymbol{x} \in \mathrm{co}S$，因为 $S \subset \mathrm{co}S$，所以 $S \cup \{\boldsymbol{x}\} \subset \mathrm{co}S$. 由此，我们有 $\mathrm{co}(S \cup \{\boldsymbol{x}\}) \subset \mathrm{co}S$. 又显然有 $\mathrm{co}S \subset \mathrm{co}(S \cup \{\boldsymbol{x}\})$，故得结论. □

下面是凸包的表示定理.

定理 1.1.8 设集合 $S \subset \mathscr{R}^n$，则对任何正整数 m 有

$$\mathrm{co}S = \left\{\boldsymbol{x} \in \mathscr{R}^n \middle| \boldsymbol{x} = \sum_{i=1}^{m} \lambda_i \boldsymbol{x}^i \quad \forall \boldsymbol{x}^i \in S,\right.$$

$$\left. \forall \lambda_i \geqslant 0 \ (i = 1, \cdots, m),\ \sum_{i=1}^{m} \lambda_i = 1\right\}, \quad (1.1.4)$$

即 $\mathrm{co}S$ 是 S 中向量的有限凸组合构成的集合.

证明 记(1.1.4)右端的集合为 D.

先证 $D \subset \mathrm{co}S$. 当 $m = 1$ 时，显然成立. 对任何正整数 $m \geqslant 2$，设 $\boldsymbol{x}^1, \cdots, \boldsymbol{x}^m \in S$，由定义 1.1.5，有 $\boldsymbol{x}^1, \cdots, \boldsymbol{x}^m \in \mathrm{co}S$. 因为 $\mathrm{co}S$ 是凸集，由定理 1.1.4 知，对任意的 $\lambda_i \geqslant 0$，$\sum_{i=1}^{m} \lambda_i = 1$，有 $\sum_{i=1}^{m} \lambda_i \boldsymbol{x}^i \in \mathrm{co}S$，所以 $D \subset \mathrm{co}S$.

现证 $\mathrm{co}S \subset D$. 任取 $\boldsymbol{x}, \boldsymbol{y} \in D$，由(1.1.4)和 D 的定义可知，存在正整数 m_1 和 m_2 以及 $\boldsymbol{x}^i \in S$，$\lambda_i \geqslant 0 \ (i = 1, \cdots, m_1)$，$\sum_{i=1}^{m_1} \lambda_i = 1$；$\boldsymbol{y}^j \in S$，$\mu_j \geqslant 0 \ (j = 1, \cdots, m_2)$，$\sum_{j=1}^{m_2} \mu_j = 1$，使得 $\boldsymbol{x} = \sum_{i=1}^{m_1} \lambda_i \boldsymbol{x}^i$ 和 $\boldsymbol{y} = \sum_{j=1}^{m_2} \mu_j \boldsymbol{y}^j$. 由此，对任意的 $\lambda \in (0, 1)$ 有

$$\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y} = \sum_{i=1}^{m_1} \lambda\lambda_i \boldsymbol{x}^i + \sum_{j=1}^{m_2} (1 - \lambda)\mu_j \boldsymbol{y}^j. \quad (1.1.5)$$

因为 $\lambda\lambda_i \geqslant 0 \ (i = 1, \cdots, m_1)$，$(1 - \lambda)\mu_j \geqslant 0 \ (j = 1, \cdots, m_2)$，

并且

$$\sum_{i=1}^{m_1}\lambda\lambda_i + \sum_{j=1}^{m_2}(1-\lambda)\mu_j = \lambda\sum_{i=1}^{m_1}\lambda_i + (1-\lambda)\sum_{j=1}^{m_2}\mu_j$$
$$= \lambda + (1-\lambda) = 1,$$

所以由(1.1.5)和 D 的定义得到 $\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in D$. 据此,由定义 1.1.1 即知 D 是凸集. 由于 $S \subset D$,而 D 是凸集,故由注 1.1.2 得到 $\mathrm{co}S \subset D$. ▯

定理 1.1.9 设集合 $S_1, S_2 \subset \mathscr{Y}$.

(1) $\mathrm{co}(S_1 \cap S_2) \subset \mathrm{co}S_1 \cap \mathrm{co}S_2$. 若 S_1 和 S_2 是凸集,则 $\mathrm{co}(S_1 \cap S_2) = \mathrm{co}S_1 \cap \mathrm{co}S_2$.

(2) $\mathrm{co}(S_1 + S_2) = \mathrm{co}S_1 + \mathrm{co}S_2$.

(3) 若 S_1 和 S_2 是凸集,则 $\mathrm{co}(S_1 \cup S_2) = \bigcup_{\lambda\in[0,1]}[\lambda S_1 + (1-\lambda)S_2]$.

证明 (1) 设 $\boldsymbol{y} \in \mathrm{co}(S_1 \cap S_2)$, 由定理 1.1.8 得知

$$\boldsymbol{y} \in \left\{\boldsymbol{x} \in \mathscr{Y} \,\middle|\, \boldsymbol{x} = \sum_{i=1}^{m}\lambda_i \boldsymbol{x}^i \quad \forall \boldsymbol{x}^i \in S_1 \cap S_2, \right.$$
$$\left. \forall \lambda_i \geqslant 0 (i = 1, \cdots, m),\ \sum_{i=1}^{m}\lambda_i = 1\right\}.$$

由此,我们有

$$\boldsymbol{y} \in \left\{\boldsymbol{x} \in \mathscr{Y} \,\middle|\, \boldsymbol{x} = \sum_{i=1}^{m}\lambda_i \boldsymbol{x}^i \quad \forall \boldsymbol{x}^i \in S_1, \right.$$
$$\left. \forall \lambda_i \geqslant 0\ (i = 1, \cdots, m),\ \sum_{i=1}^{m}\lambda_i = 1\right\}$$

和

$$\boldsymbol{y} \in \left\{\boldsymbol{x} \in \mathscr{Y} \,\middle|\, \boldsymbol{x} = \sum_{i=1}^{m}\lambda_i \boldsymbol{x}^i \quad \forall \boldsymbol{x}^i \in S_2, \right.$$
$$\left. \forall \lambda_i \geqslant 0\ (i = 1, \cdots, m),\ \sum_{i=1}^{m}\lambda_i = 1\right\},$$

从而 $\boldsymbol{y} \in \mathrm{co}S_1$ 和 $\boldsymbol{y} \in \mathrm{co}S_2$，即 $\boldsymbol{y} \in \mathrm{co}S_1 \cap \mathrm{co}S_2$.

设 S_1 和 S_2 是凸集，我们再证明 $\mathrm{co}S_1 \cap \mathrm{co}S_2 \subset \mathrm{co}(S_1 \cap S_2)$. 为此设 $\boldsymbol{y} \in \mathrm{co}S_1 \cap \mathrm{co}S_2$，因为 S_1 和 S_2 是凸的，由定理 1.1.7 的 (1) 有 $\boldsymbol{y} \in S_1 \cap S_2$. 据定义 1.1.5，便得 $\boldsymbol{y} \in S_1 \cap S_2 \subset \mathrm{co}(S_1 \cap S_2)$.

(2) 先证 $\mathrm{co}(S_1 + S_2) \subset \mathrm{co}S_1 + \mathrm{co}S_2$. 设 $\boldsymbol{y} \in \mathrm{co}(S_1 + S_2)$，由定理 1.1.8 知 $\boldsymbol{y} = \sum\limits_{i=1}^{m} \lambda_i \boldsymbol{x}^i$，其中 $\boldsymbol{x}^i \in S_1 + S_2$，$\lambda_i \geqslant 0\ (i = 1, \cdots, m)$，$\sum\limits_{i=1}^{m} \lambda_i = 1$. 于是，存在 $\boldsymbol{u}^i \in S_1$，$\boldsymbol{v}^i \in S_2$，使 $\boldsymbol{x}^i = \boldsymbol{u}^i + \boldsymbol{v}^i (i = 1, \cdots, m)$ 有

$$\boldsymbol{y} = \sum_{i=1}^{m} \lambda_i \boldsymbol{u}^i + \sum_{i=1}^{m} \lambda_i \boldsymbol{v}^i.$$

再由定理 1.1.8 得知 $\boldsymbol{y} \in \mathrm{co}S_1 + \mathrm{co}S_2$.

为证 $\mathrm{co}S_1 + \mathrm{co}S_2 \subset \mathrm{co}(S_1 + S_2)$，设 $\boldsymbol{y} \in \mathrm{co}S_1 + \mathrm{co}S_2$，则存在 $\boldsymbol{u} \in \mathrm{co}S_1$ 和 $\boldsymbol{v} \in \mathrm{co}S_2$ 使 $\boldsymbol{y} = \boldsymbol{u} + \boldsymbol{v}$. 由定理 1.1.8 我们有

$$\boldsymbol{u} = \sum_{i=1}^{m_1} \lambda_i \boldsymbol{u}^i，其中\ \boldsymbol{u}^i \in S_1，\lambda_i \geqslant 0\ (i = 1, \cdots, m_1)，\sum_{i=1}^{m_1} \lambda_i = 1，$$

$$\boldsymbol{v} = \sum_{j=1}^{m_2} \mu_j \boldsymbol{v}^j，其中\ \boldsymbol{v}^j \in S_2，\mu_j \geqslant 0\ (j = 1, \cdots, m_2)，\sum_{j=1}^{m_2} \mu_j = 1.$$

由此得到

$$\boldsymbol{y} = \sum_{i=1}^{m_1} \lambda_i \boldsymbol{u}^i + \sum_{j=1}^{m_2} \mu_j \boldsymbol{v}^j = \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \nu_{ij} (\boldsymbol{u}^i + \boldsymbol{v}^j),$$

其中 $\boldsymbol{u}^i + \boldsymbol{v}^j \in S_1 + S_2$，$\nu_{ij} = \lambda_i \mu_j \geqslant 0\ (i = 1, \cdots, m_1;\ j = 1, \cdots, m_2)$，$\sum\limits_{i=1}^{m_1} \sum\limits_{j=1}^{m_2} \nu_{ij} = 1$. 再由定理 1.1.8，即得 $\boldsymbol{y} \in \mathrm{co}(S_1 + S_2)$.

(3) 设 $\boldsymbol{y} \in \bigcup\limits_{\lambda \in [0,1]} [\lambda S_1 + (1 - \lambda) S_2]$，则存在 $\bar{\lambda} \in [0, 1]$ 使

$y \in \bar{\lambda}S_1 + (1-\bar{\lambda})S_2$. 于是，存在 $x^1 \in S_1$，$x^2 \in S_2$，有 $y \in \bar{\lambda}x^1 + (1-\bar{\lambda})x^2$. 因为 x^1，$x^2 \in S_1 \cup S_2 \subset \mathrm{co}(S_1 \cup S_2)$，而 $\mathrm{co}(S_1 \cup S_2)$ 是凸集，故得 $y \in \mathrm{co}(S_1 \cup S_2)$.

反之，设 $y \in \mathrm{co}(S_1 \cup S_2)$. 由定理 1.1.8 知，存在正整数 m 和 $y^i \in S_1 \cup S_2$，$\bar{\lambda}_i \geqslant 0$ $(i = 1, \cdots, m)$，$\sum_{i=1}^{m} \bar{\lambda}_i = 1$，使得

$$y = \sum_{i=1}^{m} \bar{\lambda}_i y^i. \tag{1.1.6}$$

若 $y^i \in S_1 (i = 1, \cdots, m)$，则因为 S_1 是凸的，由(1.1.6)有 $y \in S_1$，故 $y \in \bigcup_{\lambda \in [0,1]} [\lambda S_1 + (1-\lambda)S_2]$. 若 $y^i \in S_2 (i = 1, \cdots, m)$，则因为 S_2 是凸的，同理有 $y \in \bigcup_{\lambda \in [0,1]} [\lambda S_1 + (1-\lambda)S_2]$. 现设 $y^i \in S_1$ $(i = 1, \cdots, p)$ 和 $y^i \in S_2 (i = p+1, \cdots, m)$ $(1 \leqslant p \leqslant m-1)$. 令 $\bar{\lambda} = \sum_{i=1}^{p} \bar{\lambda}_i$，则 $\bar{\lambda} \in (0, 1)$. 由(1.1.6)我们有

$$\begin{aligned} y &= \sum_{i=1}^{p} \bar{\lambda}_i y^i + \sum_{i=p+1}^{m} \bar{\lambda}_i y^i = \bar{\lambda} \sum_{i=1}^{p} \frac{\bar{\lambda}_i}{\bar{\lambda}} y^i + (1-\bar{\lambda}) \sum_{i=p+1}^{m} \frac{\bar{\lambda}_i}{1-\bar{\lambda}} y^i \\ &= \bar{\lambda} x^1 + (1-\bar{\lambda}) x^2, \end{aligned} \tag{1.1.7}$$

其中 $x^1 = \sum_{i=1}^{p} \frac{\bar{\lambda}_i}{\bar{\lambda}} y^i$，$x^2 = \sum_{i=p+1}^{m} \frac{\bar{\lambda}_i}{1-\bar{\lambda}} y^i$. 因为 S_1 和 S_2 是凸的，故 $x^1 \in S_1$ 和 $x^2 \in S_2$. 据此，由(1.1.7)得到

$$y = \bar{\lambda} x^1 + (1-\bar{\lambda}) x^2 \in \bar{\lambda} S_1 + (1-\bar{\lambda}) S_2,$$

从而 $y \in \bigcup_{\lambda \in [0,1]} [\lambda S_1 + (1-\lambda)S_2]$. □

下面两个概念是凸包的扩展.

定义 1.1.6 设集合 $S \subset \mathscr{V}$.

(1) $\mathscr{V}$ 中所有包含 S 的仿射集的交，称为 S 的仿射包(或由 S 生成的仿射集)，记作 $\mathrm{aff} S$.

(2) $\mathscr{V}$ 中所有包含 S 的线性子空间的交，称为 S 的线性包

(或由 S 生成的线性子空间),记作 $\mathrm{lin}S$.

我们约定:空集的仿射包是空集.

注 1.1.3　由定义 1.1.6 可知,集合的仿射包是包含它的最小仿射集;集合的线性包是包含它的最小线性子空间.

定理 1.1.10　设集合 $S\subset\mathscr{V}$.

(1) 若 S 是仿射集,则 $\mathrm{aff}S=S$. 若 $\boldsymbol{x}\in\mathrm{aff}S$, 则 $\mathrm{aff}(S\cup\{\boldsymbol{x}\})=\mathrm{aff}S$.

(2) 若 S 是线性子空间,则 $\mathrm{lin}S=S$. 若 $\boldsymbol{x}\in\mathrm{lin}S$, 则 $\mathrm{lin}(S\cup\{\boldsymbol{x}\})=\mathrm{lin}S$.

证明　与定理 1.1.7 的证明类似. □

定理 1.1.11　设集合 $S\subset\mathscr{V}$.

(1) 对任何正整数 m 有

$$\mathrm{aff}S=\left\{\boldsymbol{x}\in\mathscr{V}\,\middle|\,\boldsymbol{x}=\sum_{i=1}^{m}\lambda_i\boldsymbol{x}^i\quad\forall\boldsymbol{x}^i\in S,\right.$$

$$\left.\forall\lambda_i\in R\ (i=1,\cdots,m),\ \sum_{i=1}^{m}\lambda_i=1\right\}.$$

即 $\mathrm{aff}S$ 是 S 中向量的有限仿射组合构成的集合.

(2) 对任何正整数 m 有

$$\mathrm{lin}S=\left\{\boldsymbol{x}\in\mathscr{V}\,\middle|\,\boldsymbol{x}=\sum_{i=1}^{m}\lambda_i\boldsymbol{x}^i\quad\forall\boldsymbol{x}^i\in S,\right.$$

$$\left.\forall\lambda_i\in R\ (i=1,\cdots,m)\right\}.$$

即 $\mathrm{lin}S$ 是 S 中向量的有限线性组合构成的集合.

证明　与定理 1.1.8 的证明类似. □

§1.2　凸集的结构性质

本节介绍凸集的代数结构和拓扑结构性质.先论述凸集的有关代数结构性质.

设 $\boldsymbol{x}^1$、$\boldsymbol{x}^2 \in \mathscr{Y}$，记 $(\boldsymbol{x}^1, \boldsymbol{x}^2)$、$[\boldsymbol{x}^1, \boldsymbol{x}^2]$ 以及 $(\boldsymbol{x}^1, \boldsymbol{x}^2]$ 和 $[\boldsymbol{x}^1, \boldsymbol{x}^2)$ 依次表示点 $\boldsymbol{x}^1$ 和 $\boldsymbol{x}^2$ 之间的开线段、闭线段以及半开半闭线段，即

$$
\begin{aligned}
(\boldsymbol{x}^1, \boldsymbol{x}^2) &= \{\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 \mid \boldsymbol{x}^1, \boldsymbol{x}^2 \in \mathscr{Y}, \lambda \in (0, 1)\}, \\
[\boldsymbol{x}^1, \boldsymbol{x}^2] &= \{\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 \mid \boldsymbol{x}^1, \boldsymbol{x}^2 \in \mathscr{Y}, \lambda \in [0, 1]\}, \\
(\boldsymbol{x}^1, \boldsymbol{x}^2] &= \{\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 \mid \boldsymbol{x}^1, \boldsymbol{x}^2 \in \mathscr{Y}, \lambda \in (0, 1]\}, \\
[\boldsymbol{x}^1, \boldsymbol{x}^2) &= \{\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 \mid \boldsymbol{x}^1, \boldsymbol{x}^2 \in \mathscr{Y}, \lambda \in [0, 1)\}.
\end{aligned}
\tag{1.2.1}
$$

定义 1.2.1　设集合 $S \subset \mathscr{Y}$.

(1) 集合

$S^i = \{\boldsymbol{x} \in S \mid \forall \boldsymbol{v} \in \mathscr{Y}, \exists \varepsilon > 0 \text{ 有} [\boldsymbol{x}, \boldsymbol{x} + \varepsilon \boldsymbol{v}] \subset S\}$

称为 S 的代数内部. $\boldsymbol{x} \in S^i$ 称为 S 的代数内点. 若 $S = S^i$，则称 S 是代数开集.

(2) 集合

$S^{ri} = \{\boldsymbol{x} \in S \mid \forall \boldsymbol{v} \in \text{aff}S - \boldsymbol{x}, \exists \varepsilon > 0 \text{ 有} [\boldsymbol{x}, \boldsymbol{x} + \varepsilon \boldsymbol{v}] \subset S\}$

称为 S 的相对代数内部. $\boldsymbol{x} \in S^{ri}$ 称为 S 的相对代数内点.

(3) 集合

$$S^a = \{\boldsymbol{x} \in \mathscr{Y} \mid \exists \boldsymbol{v} \in S, \text{有} [\boldsymbol{v}, \boldsymbol{x}) \subset S\}$$

称为 S 的代数闭包. 若 $S = S^a$，则称 S 是代数闭集.

(4) 集合 $S^b = S^a \backslash S^i$ 称为 S 的代数边界.

注 1.2.1　由定义 1.2.1，显然有

$$S^i \subset S^{ri} \subset S \subset S^a.$$

定理 1.2.1　设 $S \subset \mathscr{Y}$ 是凸集.

(1) 若 $\boldsymbol{x}^1 \in S^i$，$\boldsymbol{x}^2 \in S^a$，则 $[\boldsymbol{x}^1, \boldsymbol{x}^2) \subset S^i$.

(2) 若 $\boldsymbol{x}^1 \in S^{ri} \neq \varnothing$，$\boldsymbol{x}^2 \in S$，则 $[\boldsymbol{x}^1, \boldsymbol{x}^2) \subset S^{ri}$.

证明　(1) 任取 $\boldsymbol{y} \in [\boldsymbol{x}^1, \boldsymbol{x}^2)$，则有 $\lambda \in [0, 1)$ 使 $\boldsymbol{y} = \lambda \boldsymbol{x}^2 +$

$(1-\lambda)\boldsymbol{x}^1$. 以下分两种情况讨论.

(a) 设 $\boldsymbol{x}^2 \in S$. 因为 $\boldsymbol{x}^1 \in S^i \subset S$, S 是凸的, 故 $\boldsymbol{y} \in S$. 从 $\boldsymbol{x}^1 \in S^i$, 由定义 1.2.1 的(1)可知, 对任意的 $\boldsymbol{v} \in \mathscr{Y}$, 存在 $\varepsilon > 0$ 有 $[\boldsymbol{x}^1, \boldsymbol{x}^1 + \varepsilon \boldsymbol{v}] \subset S$, 从而 $\boldsymbol{x}^1 + \varepsilon \boldsymbol{v} \in S$. 由于 S 是凸的, 因此

$$\boldsymbol{y} + \lambda\varepsilon\boldsymbol{v} = \lambda\boldsymbol{x}^2 + (1-\lambda)(\boldsymbol{x}^1 + \varepsilon\boldsymbol{v}) \subset S.$$

于是, 由 $\boldsymbol{y} \in S$ 又知 $[\boldsymbol{y}, \boldsymbol{y} + \lambda\varepsilon\boldsymbol{v}] \subset S$. 因为 $\lambda\varepsilon > 0$, 再按定义 1.2.1 的(1)有 $\boldsymbol{y} \in S^i$, 由 $\boldsymbol{y}$ 的任意性得到 $[\boldsymbol{x}^1, \boldsymbol{x}^2) \subset S^i$.

(b) 设 $\boldsymbol{x}^2 \in S^a \backslash S$. 这时 $\boldsymbol{x}^2 \in S^a$, 由定义 1.2.1 的(3)可知存在 $\boldsymbol{v} \in S$ 使 $[\boldsymbol{v}, \boldsymbol{x}^2) \subset S$. 若 $\boldsymbol{x}^1$, $\boldsymbol{x}^2$, $\boldsymbol{v}$ 三点共线, 不妨设 $\boldsymbol{v} \in [\boldsymbol{x}^1, \boldsymbol{x}^2)$. 任取 $\boldsymbol{z} \in [\boldsymbol{v}, \boldsymbol{x}^2) \subset S$, 则由(a)的结果有 $[\boldsymbol{x}^1, \boldsymbol{z}) \subset S^i$. 由 $\boldsymbol{z}$ 的任意性即得 $[\boldsymbol{x}^1, \boldsymbol{x}^2) \subset S^i$. 若 $\boldsymbol{x}^1$, $\boldsymbol{x}^2$, $\boldsymbol{v}$ 不共线, 则取 $\boldsymbol{t} \in (\boldsymbol{x}^2, \boldsymbol{y})$, 从而 $\boldsymbol{y} \in (\boldsymbol{x}^1, \boldsymbol{t})$. 从 $\boldsymbol{x}^1 \in S^i$ 由定义 1.2.1 的(1)可知, 存在充分小的 $\varepsilon > 0$ 使 $\boldsymbol{u} = \boldsymbol{x}^1 + \varepsilon(\boldsymbol{x}^2 - \boldsymbol{v}) \in S$. 令 l 是过点 $\boldsymbol{u}$ 和 $\boldsymbol{t}$ 的直线, 由于 ε 充分小, 故 l 与 $(\boldsymbol{v}, \boldsymbol{x}^2)$ 交于某点 $\boldsymbol{q} \in S$. 据此, 由 S 是凸的, 得 $\boldsymbol{t} \in S$. 于是, 由(a)也推知 $[\boldsymbol{x}^1, \boldsymbol{t}) \subset S^i$, 因而 $\boldsymbol{y} \in S^i$. 由 $\boldsymbol{y}$ 的任意性即得 $[\boldsymbol{x}^1, \boldsymbol{x}^2) \subset S^i$.

(2) 任取 $\boldsymbol{y} \in [\boldsymbol{x}^1, \boldsymbol{x}^2) \subset S$, 则有 $\lambda \in [0, 1)$ 使 $\boldsymbol{y} = \lambda\boldsymbol{x}^2 + (1-\lambda)\boldsymbol{x}^1$. 从 $\boldsymbol{x}^1 \in S^{ri}$ 由定义 1.2.1 中的(2)知, 对任意的 $\boldsymbol{v} \in \mathrm{aff}S - S$ 存在 $\varepsilon > 0$ 有 $[\boldsymbol{x}^1, \boldsymbol{x}^1 + \varepsilon\boldsymbol{v}] \subset S$, 故 $\boldsymbol{x}^1 + \varepsilon\boldsymbol{v} \in S$. 因为 S 是凸的, $\boldsymbol{x}^2 \in S$, 所以

$$\boldsymbol{y} + (1-\lambda)\varepsilon\boldsymbol{v} = \lambda\boldsymbol{x}^2 + (1-\lambda)(\boldsymbol{x}^1 + \varepsilon\boldsymbol{v}) \in S.$$

从而 $[\boldsymbol{y}, \boldsymbol{y} + (1-\lambda)\varepsilon\boldsymbol{v}] \subset S$. 由此, 根据定义 1.2.1 的(2)知 $\boldsymbol{y} \in S^{ri}$. 由 $\boldsymbol{y}$ 的任意性即得结论. ▯

定理 1.2.2 设 $S \subset \mathscr{Y}$ 是凸集.

(1) S^i 是凸集.

(2) S^{ri} 是凸集.

(3) S^a 是凸集.

证明 (1) 任取 $\boldsymbol{x}^1$, $\boldsymbol{x}^2 \in S^i \subset S$ 和 $\lambda \in (0, 1)$, 因为 S 是凸

的，则 $\boldsymbol{y}=\lambda\boldsymbol{x}^1+(1-\lambda)\boldsymbol{x}^2\in S$. 由定义 1.2.1 的(1)，对任意的 $\boldsymbol{v}\in\mathscr{X}$ 存在 $\varepsilon>0$ 有

$$[\boldsymbol{x}^1,\ \boldsymbol{x}^1+\varepsilon\boldsymbol{v}]\subset S,\ [\boldsymbol{x}^2,\ \boldsymbol{x}^2+\varepsilon\boldsymbol{v}]\subset S,$$

从而 $\boldsymbol{x}^1+\varepsilon\boldsymbol{v}$，$\boldsymbol{x}^2+\varepsilon\boldsymbol{v}\in S$. 又因为 S 是凸的，故

$$\boldsymbol{y}+\varepsilon\boldsymbol{v}=\lambda(\boldsymbol{x}^1+\varepsilon\boldsymbol{v})+(1-\lambda)(\boldsymbol{x}^2+\varepsilon\boldsymbol{v})\in S.$$

据此，由 $\boldsymbol{y}\in S$ 知 $[\boldsymbol{y},\ \boldsymbol{y}+\varepsilon\boldsymbol{v}]\subset S$. 再按定义 1.2.1 的(1)得 $\boldsymbol{y}\in S^i$，故 S^i 是凸集.

(2) 与(1)的证明类似，由定义 1.2.1 的(2)可得.

(3) 任取 $\boldsymbol{x}^1$，$\boldsymbol{x}^2\in S^a$ 和 $\lambda\in(0,1)$，考虑 $\boldsymbol{y}=\lambda\boldsymbol{x}^1+(1-\lambda)\boldsymbol{x}^2$. 由定义 1.2.1 的(3)可知存在 $\boldsymbol{v}^1$，$\boldsymbol{v}^2\in S$ 使 $[\boldsymbol{v}^1,\ \boldsymbol{x}^1)\subset S$，$[\boldsymbol{v}^2,\ \boldsymbol{x}^2)\subset S$. 令 $\boldsymbol{v}=\lambda\boldsymbol{v}^1+(1-\lambda)\boldsymbol{v}^2$，则 $\boldsymbol{v}\in S$. 再设 $\mu\in(0,1)$，则有

$$\mu\boldsymbol{v}^1+(1-\mu)\boldsymbol{x}^1\in S,\ \mu\boldsymbol{v}^2+(1-\mu)\boldsymbol{x}^2\in S.$$

因为

$$\begin{aligned}&\mu\boldsymbol{v}+(1-\mu)\boldsymbol{y}\\=&\mu[\lambda\boldsymbol{v}^1+(1-\lambda)\boldsymbol{v}^2]+(1-\mu)[\lambda\boldsymbol{x}^1+(1-\lambda)\boldsymbol{x}^2]\\=&\lambda[\mu\boldsymbol{v}^1+(1-\mu)\boldsymbol{x}^1]+(1-\lambda)[\mu\boldsymbol{v}^2+(1-\mu)\boldsymbol{x}^2],\end{aligned}$$

所以 $\mu\boldsymbol{v}+(1-\mu)\boldsymbol{y}\in S$，或 $(\boldsymbol{v},\ \boldsymbol{y})\subset S$. 由此，按定义 1.2.1 的(3)得到 $\boldsymbol{y}\in S^a$，故 S^a 是凸集. □

定理 1.2.3　设 T_1，$T_2\subset\mathscr{X}$ 是非空凸集，$T_1\cap T_2=\varnothing$，$T_1\cup T_2=\mathscr{X}$.

(1) $T_1^i\cap T_2^a=\varnothing$，$T_1^a\cap T_2^i=\varnothing$.

(2) $T_1^i\cup T_2^a=\mathscr{X}$，$T_1^a\cup T_2^i=\mathscr{X}$.

(3) 若 $H=T_1^a\cap T_2^a$，则 H 是一仿射集.

(4) 若 $H=T_1^a\cap T_2^a$，$\boldsymbol{x}^1\in T_1$，$\boldsymbol{x}^2\in T_2$，则存在 $\boldsymbol{z}\in[\boldsymbol{x}^1,\ \boldsymbol{x}^2)\cap H$.

证明　(1) 先证明 $T_1^i\cap T_2^a=\varnothing$. 设若 $\boldsymbol{x}\in T_2^a\backslash T_2$，按定义 1.2.1 的(3)可知存在 $\boldsymbol{v}\in T_2$，使 $[\boldsymbol{v},\ \boldsymbol{x})\subset T_2$. 因此，由已知

$T_1 \cap T_2 = \varnothing$ 有 $[\boldsymbol{v}, \boldsymbol{x}) \cap T_1 = \varnothing$，所以 $\boldsymbol{x} \notin T_1^i$. 设若 $\boldsymbol{x} \in T_2$，因为 T_2 是凸的，则对任意的 $\boldsymbol{v} \in T_2$ 都有 $[\boldsymbol{v}, \boldsymbol{x}) \subset T_2$，故由 $T_1 \cap T_2 = \varnothing$ 得到 $\boldsymbol{x} \notin T_1^i$. 将 T_1 和 T_2 对换，即得 $T_1^a \cap T_2^i = \varnothing$.

(2) 先证 $T_1^i \cup T_2^a = \mathscr{Y}$. 设 $\boldsymbol{x} \notin T_2^a$，由定义 1.2.1 的(3)和 $T_1 \cup T_2 = \mathscr{Y}$，则对任意的 $\boldsymbol{v} \in \mathscr{Y}$ 都存在 $\boldsymbol{y} \in [\boldsymbol{v}, \boldsymbol{x}) \cap T_1$. 因为 T_1 是凸的，于是 $[\boldsymbol{x}, \boldsymbol{y}] \subset T_1$. 由定义 1.2.1 的(1)即得 $\boldsymbol{x} \in T_1^i$. 将 T_1 和 T_2 对换，即得 $T_1^a \cup T_2^i = \mathscr{Y}$.

(3) 由 T_1、T_2 是凸的，根据定理 1.2.2 的(3)知 T_1^a，T_2^a 是凸的. 因此，由定理 1.1.2 的(2)知 $H = T_1^a \cap T_2^a$ 是凸集，故设 $\boldsymbol{x}^1$，$\boldsymbol{x}^2 \in H$，则有 $[\boldsymbol{x}^1, \boldsymbol{x}^2] \subset H$. 为证 H 是仿射集，反之假设存在 $\lambda \in R$，使 $\boldsymbol{y} = \lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 \notin H$. 因为 $[\boldsymbol{x}^1, \boldsymbol{x}^2] \subset H$，则 $\boldsymbol{y} \notin [\boldsymbol{x}^1, \boldsymbol{x}^2]$，于是有 $\boldsymbol{x}^1 \in (\boldsymbol{y}, \boldsymbol{x}^2)$ 或 $\boldsymbol{x}^2 \in (\boldsymbol{x}^1, \boldsymbol{y})$. 不妨设 $\boldsymbol{x}^1 \in (\boldsymbol{y}, \boldsymbol{x}^2)$，则由 $\boldsymbol{y} \notin H$，有 $\boldsymbol{y} \in T_1^i$ 或 $\boldsymbol{y} \in T_2^i$. 再不妨设 $\boldsymbol{y} \in T_1^i$，则由定理 1.2.1 的(1)得 $[\boldsymbol{y}, \boldsymbol{x}^2) \subset T_1^i$，从而 $\boldsymbol{x}^1 \in T_1^i$. 据此，由(1)推得 $\boldsymbol{x}^1 \notin T_2^a$，导致与 $\boldsymbol{x}^1 \in H$ 矛盾. 于是，按定义 1.1.3 的(1)得知 H 是仿射集.

(4) 先设 $\boldsymbol{x}^1 \notin T_1^i$. 由已证得的(2)推知 $\boldsymbol{x}^1 \in T_2^a$，再由 $\boldsymbol{x}^1 \in T_1 \subset T_1^a$，得 $\boldsymbol{x}^1 \in T_1^a \cap T_2^a = H$. 因此，存在 $\boldsymbol{z} = \boldsymbol{x}^1 \in [\boldsymbol{x}^1, \boldsymbol{x}^2) \cap H$. 再设 $\boldsymbol{x}^1 \in T_1^i$. 反之假设 $[\boldsymbol{x}^1, \boldsymbol{x}^2) \cap H \neq \varnothing$，则对任意的 $\boldsymbol{z} \in [\boldsymbol{x}^1, \boldsymbol{x}^2)$，有 $\boldsymbol{z} \notin H = T_1^a \cap T_2^a$，从而 $\boldsymbol{z} \notin T_1^a$ 或 $\boldsymbol{z} \notin T_2^a$. 不妨设 $\boldsymbol{z} \notin T_1^a$，则由(2)得 $\boldsymbol{z} \in T_2^i$. 因为 $\boldsymbol{x}^2 \in T_2^i$，$T_2$ 是凸的，故 $(\boldsymbol{x}^2, \boldsymbol{z}) \subset T_2$，由 $\boldsymbol{z} \in [\boldsymbol{x}^1, \boldsymbol{x}^2)$ 的任意性有 $(\boldsymbol{x}^2, \boldsymbol{x}^1) \subset T_2$，于是 $[\boldsymbol{x}^2, \boldsymbol{x}^1) \subset T_2$. 据此，由定义 1.2.1 的(3)得 $\boldsymbol{x}^1 \in T_2^a$，因而存在 $\boldsymbol{x}^1 \in T_1^i \cap T_2^a$，这导致与已知的(1)矛盾. □

以下讨论凸集和凸包的拓扑结构性质. 设 $\mathscr{X}$ 是线性拓扑空间.

定义 1.2.2 设集合 $S \subset \mathscr{X}$. 记点 $\boldsymbol{x} \in \mathscr{X}$ 的邻域为 $U(\boldsymbol{x})$.

(1) S 的所有开子集的并称为 S 的内部，记作 $\mathrm{int}S$，即

$$\mathrm{int}S = \{\boldsymbol{x} \in S \mid \exists U(\boldsymbol{x}) \text{ 使 } U(\boldsymbol{x}) \subset S\}.$$

$\boldsymbol{x} \in \mathrm{int}S$ 称为是 S 的内点. $S = \mathrm{int}S$ 意味着 S 是开集.

(2) 集合

$$\mathrm{ri}S = \{x \in S \mid \exists U(x) \text{ 使 } U(x) \cap \mathrm{aff}S \subset S\}$$

称为 S 的相对内部. $x \in \mathrm{ri}S$ 称为 S 的相对内点. 若 $S = \mathrm{ri}S$，则称 S 是相对开集.

(3) 所有包含 S 的闭集的交称为 S 的闭包，记作 $\mathrm{cl}S$ 或 $\overline{S}$. $x \in \mathrm{cl}S$ 称为是 S 的接触点. $S = \mathrm{cl}S$ 意味着 S 是闭集.

(4) 集合 $\partial S = \mathrm{cl}S \backslash \mathrm{int}S$ 称为 S 的边界. 集合 $\mathrm{rb}S = \mathrm{cl}S \backslash \mathrm{ri}S$ 称为 S 的相对边界.

注 1.2.2　设 $S \subset \mathscr{X}$，由定义 1.2.2，显然有

$$\mathrm{int}S \subset \mathrm{ri}S \subset S \subset \mathrm{cl}S.$$

又由定义 1.2.1 和定义 1.2.2，可以推知

$$\mathrm{int}S \subset S^i,\ S^a \subset \mathrm{cl}S.$$

定理 1.2.4　设 $S \subset \mathscr{X}$ 是非空凸集，则 $\mathrm{ri}S = \mathrm{int}S$ 当且仅当 $\mathrm{int}S \neq \varnothing$.

证明　充分性. 设 $\mathrm{int}S \neq \varnothing$，则 S 是无限集，从而 $\mathrm{aff}S = \mathscr{X}$. 据此，由定义 1.2.2 的(1)和(2)，我们有

$$\begin{aligned}\mathrm{ri}S &= \{x \in S \mid \exists U(x) \text{ 使 } U(x) \cap \mathrm{aff}S \subset S\} \\ &= \{x \in S \mid \exists U(x) \text{ 使 } U(x) \subset S\} = \mathrm{int}S.\end{aligned}$$

必要性. 由 S 是凸的，可以推知 $\mathrm{ri}S \neq \varnothing$. 因此，$\mathrm{int}S = \mathrm{ri}S \neq \varnothing$. □

定理 1.2.5　设 $S \subset \mathscr{X}$ 是非空凸集.

(1) 若 $x^1 \in \mathrm{int}S$，$x^2 \in \mathrm{cl}S$，则 $[x^1, x^2) \subset \mathrm{int}S$.

(2) 若 $x^1 \in \mathrm{ri}S$，$x^2 \in \mathrm{cl}S$，则 $[x^1, x^2) \subset \mathrm{ri}S$.

证明　(1) 任取 $y \in [x^1, x^2)$，则有 $\lambda \in [0, 1)$ 使 $y = \lambda x^2 + (1-\lambda)x^1$. 从 $x^1 \in \mathrm{int}S$ 由定义 1.2.2 的(1)可知，存在 x^1 的邻域 $U(x^1) \subset S$. 令 $W = \dfrac{1}{\lambda}[y - (1-\lambda)U(x^1)]$，则它是 x^2 的一个邻

域. 由 $x^2 \in \mathrm{cl}S$ 知,存在 $u^2 \in W \cap S$. 设 $u^2 = \frac{1}{\lambda}[y - (1-\lambda)u^1]$,其中 $u^1 \in U(x^1)$,则 $y = \lambda u^2 + (1-\lambda)u^1$. 因为 $U(x^1)$ 是 u^1 的邻域,故 $U = \lambda u^2 + (1-\lambda)U(x^1)$ 是 y 的邻域. 由于 $u^2 \in S$, $U(x^1) \subset S$, S 是凸集,因此得知 $U \subset S$,从而 $y \in \mathrm{int}S$.

(2)与(1)的证明类似,由定义 1.2.2 的(2)可证明. ▯

定理 1.2.6 设 $S \subset \mathscr{X}$ 是凸集.

(1) $\mathrm{int}S$ 是凸集.

(2) $\mathrm{ri}S$ 是凸集.

(3) $\mathrm{cl}S$ 是凸集.

证明 (1) 若 $\mathrm{int}S = \varnothing$,则 $\mathrm{int}S$ 是凸的. 设 $\mathrm{int}S \neq \varnothing$,任取 $x^1, x^2 \in \mathrm{int}S$,则 $x^1 \in \mathrm{int}S \subset \mathrm{cl}S$, $x^2 \in \mathrm{int}S$. 由定理 1.2.5 的(1),有 $(x^1, x^2] \in \mathrm{int}S$,或按(1.2.1)知,对任意的 $\lambda \in (0, 1)$ 有 $\lambda x^1 + (1-\lambda)x^2 \in \mathrm{int}S$. 据此,由定义 1.1.1 便知 $\mathrm{int}S$ 是凸的.

(2) 若 $\mathrm{ri}S = \varnothing$,则 $\mathrm{ri}S$ 是凸的. 设 $\mathrm{ri}S \neq \varnothing$,任取 $x^1, x^2 \in \mathrm{ri}S$,则 $x^1 \in \mathrm{ri}S \subset \mathrm{cl}S$, $x^2 \in \mathrm{ri}S$. 由定理 1.2.5 的(2),类似于(1)的证明可得证.

(3) 任取 $x^1, x^2 \in \mathrm{cl}S$ 和任意的 $\lambda \in (0, 1)$,考虑 $y = \lambda x^1 + (1-\lambda)x^2$ 的邻域 $U(y)$. 由于集值映射 $U_y: \mathscr{X} \times \mathscr{X} \to 2^{\mathscr{X}}$, $(x^1, x^2) \mapsto \lambda x^1 + (1-\lambda)x^2$ 是连续的,故存在 x^1 的邻域 $U(x^1)$ 和 x^2 的邻域 $U(x^2)$,使 $\lambda U(x^1) + (1-\lambda)U(x^2) \subset U(y)$. 另外,从 $x^1, x^2 \in \mathrm{cl}S$ 知存在 $u^1 \in U(x^1) \cap S$, $u^2 \in U(x^2) \cap S$,于是存在

$$\lambda u^1 + (1-\lambda)u^2 \in [\lambda U(x^1) + (1-\lambda)U(x^2)] \cap S$$

$$\subset U(y) \cap S.$$

因为 $U(y) \cap S \neq \varnothing$,所以 $y \in \mathrm{cl}S$. ▯

定理 1.2.7 设 $S \subset \mathscr{X}$ 是凸集.

(1) $\mathrm{cl}S = \mathrm{cl}(\mathrm{cl}S) = \mathrm{cl}(\mathrm{ri}S)$.

(2) $\mathrm{ri}S = \mathrm{ri}(\mathrm{cl}S) = \mathrm{ri}(\mathrm{ri}S)$.

(3) $\mathrm{rb}S = \mathrm{rb}(\mathrm{cl}S) = \mathrm{rb}(\mathrm{ri}S)$.

(4) $\mathrm{aff}S=\mathrm{aff}(\mathrm{cl}S)=\mathrm{aff}(\mathrm{ri}S)$.

证明　由定义 1.2.2,利用定理 1.2.5 可以推证. □

定理 1.2.8　设 $S\subset\mathscr{X}$ 是凸集.

(1) 若 $\mathrm{int}S\neq\varnothing$, 则 $\mathrm{int}S=S^{i}$, $\mathrm{cl}S=S^{a}$.

(2) 若 $\mathrm{ri}S\neq\varnothing$, 则 $\mathrm{ri}S=S^{ri}$, $\mathrm{cl}S=S^{a}$.

证明　(1) 先证 $\mathrm{int}S=S^{i}$. 由注 1.2.2 有 $\mathrm{int}S\subset S^{i}$, 以下证 $S^{i}\subset\mathrm{int}S$. 为此,任取 $\boldsymbol{x}\in S^{i}$, 设 $\boldsymbol{v}\in\mathrm{int}S$, 则存在 $\boldsymbol{y}$ 使 $\boldsymbol{x}\in(\boldsymbol{y},\boldsymbol{v})\subset S$, 故 $\boldsymbol{y}\in\mathrm{cl}S$. 由定理 1.2.5 的(1)知 $(\boldsymbol{y},\boldsymbol{v})\subset\mathrm{int}S$, 因为 $\boldsymbol{x}\in(\boldsymbol{y},\boldsymbol{v})$, 于是得到 $\boldsymbol{x}\in\mathrm{int}S$.

以下证 $\mathrm{cl}S=S^{a}$. 由注 1.2.2,有 $S^{a}\subset\mathrm{cl}S$. 为证 $\mathrm{cl}S\subset S^{a}$, 任取 $\boldsymbol{x}\in\mathrm{cl}S$, 设 $\boldsymbol{v}\in\mathrm{int}S$, 由定理 1.2.5 的(1)知 $[\boldsymbol{v},\boldsymbol{x})\subset\mathrm{int}S$, 从而 $[\boldsymbol{v},\boldsymbol{x})\subset S$. 据此,由定义 1.2.1 的(3)得 $\boldsymbol{x}\in S^{a}$.

(2)与(1)的证明类似,由定义 1.2.1 和定义 1.2.2,利用定理 1.2.5 中的(2)可以推证. □

定理 1.2.9　设 $S\subset\mathscr{X}$ 是开集,则 $\mathrm{co}S$ 是开集.

证明　由 S 是开集,则 $S=\mathrm{int}S$. 因为 $S\subset\mathrm{co}S$, 故 $S=\mathrm{int}S\subset\mathrm{int}(\mathrm{co}S)$. 由于 $\mathrm{co}S$ 是凸的,根据定理 1.2.6 的(1)知 $\mathrm{int}(\mathrm{co}S)$ 是凸集,于是由注 1.1.2 有 $\mathrm{co}S\subset\mathrm{int}(\mathrm{co}S)$. 又显然有 $\mathrm{int}(\mathrm{co}S)\subset\mathrm{co}S$, 因此 $\mathrm{co}S=\mathrm{int}(\mathrm{co}S)$, 故 $\mathrm{co}S$ 是开集. □

定义 1.2.3　设集合 $S\subset\mathscr{X}$. $\mathscr{X}$ 中所有包含 S 的闭凸集的交称为是 S 的闭凸包,记作 $\overline{\mathrm{co}}S$.

注 1.2.3　由定义 1.2.3、定义 1.2.2 的(3)和定理 1.1.7 的(1),若 S 是凸集,则 $\overline{\mathrm{co}}S=\overline{\mathrm{co}S}=\overline{S}$.

定理 1.2.10　设集合 $S\subset\mathscr{X}$ 非空,则 $\overline{\mathrm{co}}S=\mathrm{cl}(\mathrm{co}S)$.

证明　因为 $\mathrm{co}S$ 是包含 S 的最小凸集,而 $\overline{\mathrm{co}}S$ 是包含 S 的闭凸集,故 $\mathrm{co}S\subset\overline{\mathrm{co}}S$, 从而 $\mathrm{cl}(\mathrm{co}S)\subset\overline{\mathrm{co}}S$. 反之,由 $\overline{\mathrm{co}}S$ 是包含 S 的最小闭凸集,又有 $\overline{\mathrm{co}}S\subset\mathrm{cl}(\mathrm{co}S)$. 而得知 $\overline{\mathrm{co}}S=\mathrm{cl}(\mathrm{co}S)$. □

注 1.2.4　闭凸包并不是闭集的凸包. 事实上,闭集的凸包可以不是闭集. 而定理 1.2.10 表明,闭凸包与凸包的闭包是相同的.

最后,给出线性拓扑空间中严格凸集的概念.

定义 1.2.4 设集合 $S\subset\mathscr{X}$. 若对任意的 $\lambda\in(0,1)$ 有

$$\lambda \boldsymbol{x}^1+(1-\lambda)\boldsymbol{x}^2\in \operatorname{int}S\quad \forall \boldsymbol{x}^1,\boldsymbol{x}^2\in S,\ \boldsymbol{x}^1\neq \boldsymbol{x}^2,$$

则称 S 是($\mathscr{X}$ 中的)**严格凸集**,或集合 S 是**严格凸的**.

注 1.2.5 由定义 1.2.4 和定义 1.1.1 易知,若 S 是严格凸集,则 S 显然也是凸集,但反之不然.

§1.3 凸集分离定理

这一节讨论在线性空间和线性拓扑空间中两个凸集被超平面分离的问题. 本节中介绍的凸集分离定理是凸分析和非光滑分析的重要分析工具.

先介绍几个有关的概念和引理.

设 $\mathscr{Y}$ 是线性空间, $\mathscr{L}\subset\mathscr{Y}$ 是一线性子空间. 若存在 $\boldsymbol{v}\in\mathscr{Y}\backslash\mathscr{L}$, 使得

$$\mathscr{Y}=\mathscr{L}+R\boldsymbol{v}=\{\boldsymbol{a}+\lambda\boldsymbol{v}\mid \boldsymbol{a}\in\mathscr{L},\ \lambda\in R\},$$

则称 $\mathscr{L}$ 具有余维数 1.

定义 1.3.1 设 $\mathscr{L}\subset\mathscr{Y}$ 是具有余维数 1 的线性子空间, $\boldsymbol{b}\in\mathscr{Y}$, 则称仿射集 $H=\mathscr{L}+\boldsymbol{b}$ 是 $\mathscr{Y}$ 中的**超平面**.

设 $\mathscr{Y}^*$ 是 $\mathscr{Y}$ 的对偶空间,即它是 $\mathscr{Y}$ 上线性泛函全体构成的空间. $\langle\boldsymbol{x}^*,\boldsymbol{x}\rangle$表示泛函 $\boldsymbol{x}^*\in\mathscr{Y}^*$ 在点 $\boldsymbol{x}\in\mathscr{Y}$ 处的值.

引理 1.3.1 设集合 $H\subset\mathscr{Y}$ 非空,则 H 是超平面当且仅当存在非零实线性泛函 $\boldsymbol{x}^*\in\mathscr{Y}^*$ 和实数 $\beta\in R$,使得

$$H=\{\boldsymbol{x}\in\mathscr{Y}\mid\langle\boldsymbol{x}^*,\boldsymbol{x}\rangle=\beta\}. \tag{1.3.1}$$

证明 充分性. 设存在 $\boldsymbol{x}^*\in\mathscr{Y}^*\backslash\{\boldsymbol{0}\}$ 和 $\beta\in R$ 使得(1.3.1)成立. 因为 $\boldsymbol{x}^*$ 是实线性泛函,所以 $\mathscr{L}=\{\boldsymbol{x}\in\mathscr{Y}\mid\langle\boldsymbol{x}^*,\boldsymbol{x}\rangle=0\}$是 $\mathscr{Y}$ 中的线性子空间. 又 $\boldsymbol{x}^*$ 是实线性泛函,故存在 $\boldsymbol{v}\in\mathscr{Y}$ 使 $\langle\boldsymbol{x}^*,\boldsymbol{v}\rangle=1$. 于是, 对任意的 $\boldsymbol{x}\in\mathscr{Y}$ 有 $\langle\boldsymbol{x}^*,\boldsymbol{x}-\langle\boldsymbol{x}^*,\boldsymbol{x}\rangle\boldsymbol{v}\rangle=\langle\boldsymbol{x}^*,$

$\boldsymbol{x}\rangle - \langle \boldsymbol{x}^*, \boldsymbol{x}\rangle = 0$，因而 $\boldsymbol{x} - \langle \boldsymbol{x}^*, \boldsymbol{x}\rangle \boldsymbol{v} \in \mathscr{L}$. 据此，有 $\boldsymbol{x} \in \mathscr{L} + \langle \boldsymbol{x}^*, \boldsymbol{x}\rangle \boldsymbol{v}$，故 $\mathscr{V} \subset \mathscr{L} + R\boldsymbol{v}$. 另外，显然有 $\mathscr{L} + R\boldsymbol{v} \subset \mathscr{V}$，因此得 $\mathscr{V} = \mathscr{L} + R\boldsymbol{v}$. 由于 $\mathscr{L}$ 是具有余维数 1 的线性子空间，所以由定义 1.3.1 知，$H = \mathscr{L} + \beta\boldsymbol{v}$ 是 $\mathscr{V}$ 中的超平面.

必要性. 设 $H = \mathscr{L} + \boldsymbol{b}$ 是 $\mathscr{V}$ 中的超平面，其中 $\mathscr{L} \subset \mathscr{V}$ 是具有余维数 1 的线性子空间，$\boldsymbol{b} \in \mathscr{V}$. 取 $\boldsymbol{v} \in \mathscr{V} \backslash \mathscr{L}$，则对任一 $\boldsymbol{x} \in \mathscr{V}$ 有 $\boldsymbol{x} = \boldsymbol{a} + \lambda\boldsymbol{v}$，其中 $\boldsymbol{a} \in \mathscr{L}$，$\lambda \in R$. 设 $\boldsymbol{x}^* \in \mathscr{V}^* \backslash \{\boldsymbol{0}\}$，令 $\langle \boldsymbol{x}^*, \boldsymbol{x}\rangle = \lambda$，并且 $\langle \boldsymbol{x}^*, \boldsymbol{v}\rangle = 1$，当 $\boldsymbol{x} \in \mathscr{L}$ 时 $\langle \boldsymbol{x}^*, \boldsymbol{x}\rangle = 0$. 再令 $\beta = \langle \boldsymbol{x}^*, \boldsymbol{b}\rangle$，则显然有

$$H = \mathscr{L} + \boldsymbol{b} \subset \{\boldsymbol{x} \in \mathscr{V} \mid \langle \boldsymbol{x}^*, \boldsymbol{x}\rangle = \beta\}.$$

另外，可以证明 $\{\boldsymbol{x} \in \mathscr{V} \mid \langle \boldsymbol{x}^*, \boldsymbol{x}\rangle = \beta\} \subset H$，因此得到 $H = \{\boldsymbol{x} \in \mathscr{V} \mid \langle \boldsymbol{x}^*, \boldsymbol{x}\rangle = \beta\}$. □

引理 1.3.2 设 $T_1, T_2 \subset \mathscr{V}$ 是非空凸集，$T_1 \cap T_2 = \varnothing$，$T_1 \cup T_2 = \mathscr{V}$. 若 $H = T_1^a \cap T_2^a$，并且 $H \neq \mathscr{V}$，则 H 是一超平面.

证明 设 $\boldsymbol{h} \in H$，因为 $H = (H - \boldsymbol{h}) + \boldsymbol{h}$，由定理 1.2.3 的(3)知 $H = T_1^a \cap T_2^a$ 是仿射集. 因此，只要证明线性子空间 $H - \boldsymbol{h}$ 具有余维数 1，由定义 1.3.1 即得 H 是一超平面.

设 $\boldsymbol{b} \notin H = T_1^a \cap T_2^a$，根据定理 1.2.3 的(1)可知，若 $\boldsymbol{b} \notin T_2^a$，则 $\boldsymbol{b} \in T_1^i$；若 $\boldsymbol{b} \notin T_1^a$，则 $\boldsymbol{b} \in T_2^i$. 于是有 $\boldsymbol{b} \in T_1^i \cup T_2^i$. 不妨设 $\boldsymbol{b} \in T_1^i$，则有 $2\boldsymbol{h} - \boldsymbol{b} \notin H$（否则，将有 $\boldsymbol{b} \in H$，这与已设矛盾），因此 $2\boldsymbol{h} - \boldsymbol{b} \in T_1^i \cup T_2^i$. 用反证法可得 $2\boldsymbol{h} - \boldsymbol{b} \notin T_1^i$（因为若 $2\boldsymbol{h} - \boldsymbol{b} \in T_1^i$，由 $\boldsymbol{b} \in T_1^i$ 和 T_1^i 是凸的，则有 $\boldsymbol{h} \in T_1^i$，导致矛盾），从而知 $2\boldsymbol{h} - \boldsymbol{b} \in T_2^i$. 现在任取 $\boldsymbol{x} \in \mathscr{V}$，若 $\boldsymbol{x} + \boldsymbol{h} \in T_1$，则 $[\boldsymbol{x} + \boldsymbol{h}, 2\boldsymbol{h} - \boldsymbol{b}) \cap H \neq \varnothing$，故存在 $\boldsymbol{v} \in H$ 和 $\lambda \in (0, 1]$ 使

$$\boldsymbol{v} = \lambda(\boldsymbol{x} + \boldsymbol{h}) + (1 - \lambda)(2\boldsymbol{h} - \boldsymbol{b}).$$

由此，我们有

$$\boldsymbol{x} = \left(\frac{1}{\lambda} - 1\right)(\boldsymbol{b} - \boldsymbol{h}) + \frac{1}{\lambda}(\boldsymbol{v} - h).$$

若 $\boldsymbol{x}+\boldsymbol{h}\in T_2$，则 $[\boldsymbol{x}+\boldsymbol{h},\boldsymbol{b})\cap H\neq\varnothing$，故存在 $\boldsymbol{v}\in H$ 和 $\lambda\in(0,1]$ 使

$$\boldsymbol{v}=\lambda(\boldsymbol{x}+\boldsymbol{h})+(1-\lambda)\boldsymbol{b},$$

从而

$$\boldsymbol{x}=\left(1-\frac{1}{\lambda}\right)(\boldsymbol{b}-\boldsymbol{h})+\frac{1}{\lambda}(\boldsymbol{v}-\boldsymbol{h}).$$

综合以上讨论，对任意的 $\boldsymbol{x}\in\mathscr{V}$，总存在 $\boldsymbol{y}\in\boldsymbol{H}-\boldsymbol{h}$ 和 $\mu\in R$ 使得

$$\boldsymbol{x}=\boldsymbol{y}+\mu(\boldsymbol{b}-\boldsymbol{h})\subset H-\boldsymbol{h}+R(\boldsymbol{b}-\boldsymbol{h}).$$

据此，得到 $\mathscr{V}=H-\boldsymbol{h}+R(\boldsymbol{b}-\boldsymbol{h})$，这表明线性子空间 $H-\boldsymbol{h}$ 具有余维数 1. 于是引理得证. ▯

引理 1.3.3 设 $S_1,S_2\subset\mathscr{V}$ 是凸集. 若 $S_1\cap S_2=\varnothing$，则存在凸集 $T_1,T_2\subset\mathscr{V}$，使得 $S_1\subset T_1$，$S_2\subset T_2$，并且 $T_1\cap T_2=\varnothing$，$T_1\cup T_2=\mathscr{V}$.

证明 设 $P,Q\subset\mathscr{V}$ 是凸集. 作集合

$$\mathscr{M}=\{(P,Q)\subset\mathscr{V}\times\mathscr{V}\mid S_1\subset P,\ S_2\subset Q,\ P\cap Q=\varnothing\},$$

并且定义 $\mathscr{M}$ 中的序关系 $\precsim$ 如下：

$$(P_1,Q_1)\precsim(P_2,Q_2)\Longleftrightarrow P_1\subset P_2,\ Q_1\subset Q_2.$$

由于 $\mathscr{M}$ 的每一全序子集 $\{(P_d,Q_d)\}$ 在 $\mathscr{M}$ 中都有一个上界，即 $\left(\bigcup\limits_d P_d,\ \bigcup\limits_d Q_d\right)$，并且显然有 $S_1\subset\bigcup\limits_d P_d$，$S_2\subset\bigcup\limits_d Q_d$，因而根据 Zorn 引理[11]，$\mathscr{M}$ 必有极大元 (T_1,T_2)，并且 $S_1\subset T_1$，$S_2\subset T_2$，$T_1\cap T_2=\varnothing$. 现在再证明有 $T_1\cup T_2=\mathscr{V}$. 事实上，假若 $T_1\cup T_2\neq\mathscr{V}$，则存在 $\boldsymbol{x}\in\mathscr{V}$ 有 $\boldsymbol{x}\notin T_1\cup T_2$. 于是，由 $T_1\cap T_2=\varnothing$ 可以推知 $\mathrm{co}(\{\boldsymbol{x}\}\cup T_1)\cap T_2=\varnothing$ 或 $\mathrm{co}(\{\boldsymbol{x}\}\cup T_2)\cap T_1=\varnothing$. 不妨设有 $\mathrm{co}(\{\boldsymbol{x}\}\cup T_1)\cap T_2=\varnothing$，记 $T_1'=\mathrm{co}(\{\boldsymbol{x}\}\cup T_1)$，则 T_1' 是凸集. 由 $S_1\subset T_1\subset T_1'$ 和 $S_2\subset T_2$ 得知 $(T_1',T_2)\subset\mathscr{M}$，这表明 (T_1,T_2) 不是极大元，导致与已知矛盾. ▯

设有超平面(1.3.1)，我们称集合

$$H_-^c = \{x \in \mathscr{Y} \mid \langle x^*, x\rangle \leqslant \beta\}$$

和

$$H_+^c = \{x \in \mathscr{Y} \mid \langle x^*, x\rangle \geqslant \beta\}$$

是由超平面 H 确定的闭半空间；称集合

$$H_-^o = \{x \in \mathscr{Y} \mid \langle x^*, x\rangle < \beta\}$$

和

$$H_+^o = \{x \in \mathscr{Y} \mid \langle x^*, x\rangle > \beta\}$$

是由超平面 H 确定的开半空间. 显然，上述 4 个半空间构成两组互补的凸集，即

$$H_-^c \cup H_+^o = \mathscr{Y}, \quad H_-^c \cap H_+^o = \varnothing,$$

$$H_-^o \cup H_+^c = \mathscr{Y}, \quad H_-^o \cap H_+^c = \varnothing.$$

下面讨论两个集合被超平面分离的问题.

定义 1.3.2　设集合 $S_1, S_2 \subset \mathscr{Y}$ 非空.

(1) 若存在超平面 $H \subset \mathscr{Y}$ 使 $S_1 \subset H_-^c$ 和 $S_2 \subset H_+^c$，即存在 $x^* \in \mathscr{Y}^* \setminus \{\mathbf{0}\}$ 和 $\beta \in R$ 使

$$\langle x^*, x^1\rangle \leqslant \beta \leqslant \langle x^*, x^2\rangle \quad \forall x^1 \in S_1, \forall x^2 \in S_2, \tag{1.3.2}$$

则称集合 S_1 和 S_2 被超平面 H 分离，或 H 分离 S_1 和 S_2. 若再有 $S_1 \cup S_2 \not\subset H$，则称 H 正常分离 S_1 和 S_2，否则，称 H 非正常分离 S_1 和 S_2.

(2) 若存在超平面 $H \subset \mathscr{Y}$ 使 $S_1 \subset H_-^c$ 和 $S_2 \subset H_+^o$ 或 $S_1 \subset H_-^o$ 和 $S_2 \subset H_+^c$，即存在 $x^* \in \mathscr{Y}^* \setminus \{\mathbf{0}\}$ 和 $\beta \in R$ 使

$$\langle x^*, x^1\rangle \leqslant \beta < \langle x^*, x^2\rangle \quad \forall x^1 \in S_1, \forall x^2 \in S_2, \tag{1.3.3}$$

或

$$\langle \boldsymbol{x}^*, \boldsymbol{x}^1 \rangle < \beta \leqslant \langle \boldsymbol{x}^*, \boldsymbol{x}^2 \rangle \quad \forall \boldsymbol{x}^1 \in S_1, \forall \boldsymbol{x}^2 \in S_2. \tag{1.3.4}$$

则称集合 S_1 和 S_2 被超平面 H 严格分离，或 H 严格分离 S_1 和 S_2.

(3) 若存在 $\boldsymbol{x}^* \in \mathscr{Y}^* \backslash \{\boldsymbol{0}\}$ 和 $\beta \in R$，使

$$\sup_{x^1 \in S_1} \langle \boldsymbol{x}^*, \boldsymbol{x}^1 \rangle \leqslant (<) \beta < (\leqslant) \inf_{x^2 \in S_2} \langle \boldsymbol{x}^*, \boldsymbol{x}^2 \rangle, \tag{1.3.5}$$

则称集合 S_1 和 S_2 被超平面 H 强分离，或 H 强分离 S_1 和 S_2.

注 1.3.1 显然，若超平面 H 强分离集合 S_1 和 S_2，则它必严格分离 S_1 和 S_2；若 H 严格分离 S_1 和 S_2，则它分离 S_1 和 S_2；反之不然.

图 1.3.1 依次是正常分离，非正常分离，严格分离和强分离的情况.

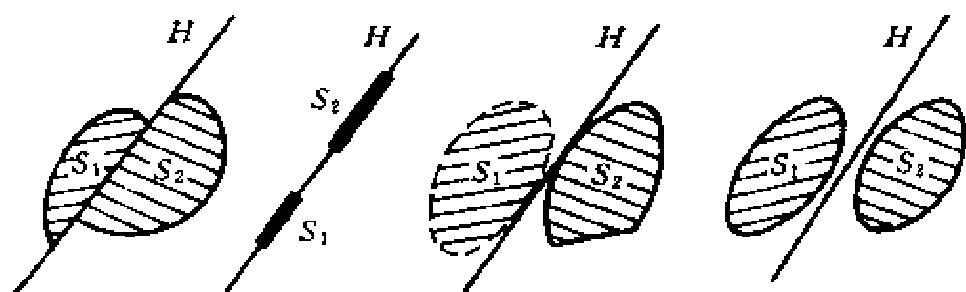


图 1.3.1

下面我们给出线性空间 $\mathscr{Y}$ 中的几个凸集分离定理.

定理 1.3.4 设 S_1、$S_2 \subset \mathscr{Y}$ 是凸集，$S_1 \neq \varnothing$，$S_2^i \neq \varnothing$，并且 $S_1 \cap S_2^i = \varnothing$.

(1) 存在超平面 $H \subset \mathscr{Y}$ 正常分离 S_1 和 S_2.

(2) 存在超平面 $H \subset \mathscr{Y}$ 严格分离 S_1 和 S_2^i.

证明 (1) 因为 S_2 是凸的，由定理 1.2.2 的(1)，则 S_2^i 是凸的. 又由 S_1 是凸的和 $S_1 \cap S_2^i = \varnothing$，根据引理 1.3.3 可知，存在凸集 T_1，$T_2 \subset \mathscr{Y}$ 使 $S_1 \subset T_1$，$S_2^i \subset T_2$，并且 $T_1 \cap T_2 = \varnothing$，$T_1 \cup T_2 = \mathscr{Y}$. 从 S_1 和 S_2^i 是非空凸集可知 T_1 和 T_2 是非空凸集，于是由

定理 1.2.3 的(1)得

$$T_1^a \cap T_2^i = \varnothing. \tag{1.3.6}$$

另外,由 $S_2^i \subset T_2$ 有 $(S_2^i)^i \subset T_2^i$,从定义 1.2.1 的(1)和定理 1.2.1 的(1)可推知 $(S_2^i)^i = S_2^i$,故 $S_2^i \subset T_2^i$. 因此,由(1.3.6)我们有

$$T_1^a \cap S_2^i = \varnothing. \tag{1.3.7}$$

设 $H = T_1^a \cap T_2^a$,因为 $S_2^i \neq \varnothing$,故存在 $\boldsymbol{y} \in S_2^i \subset \mathscr{Y}$,由(1.3.7)知 $\boldsymbol{y} \notin T_1^a$. 于是有 $\boldsymbol{y} \notin T_1^a \cap T_2^a = H$,从而 $H \neq \mathscr{Y}$. 据此,由引理 1.3.2 得知 H 是一超平面. 再由引理 1.3.1,可设

$$H = \{\boldsymbol{x} \in \mathscr{Y} \mid \langle \boldsymbol{x}^*, \boldsymbol{x} \rangle = \beta\}, \tag{1.3.8}$$

其中 $\boldsymbol{x}^* \in \mathscr{Y}^* \setminus \{\boldsymbol{0}\}$, $\beta \in R$. 因为据(1.3.7)有

$$H \cap S_2^i = T_1^a \cap T_2^a \cap S_2^i = \varnothing, \tag{1.3.9}$$

所以对任意的 $\boldsymbol{x} \in S_2^i$,由(1.3.8)有 $\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle \neq \beta$. 不妨设存在 $\boldsymbol{v} \in S_2^i$ 使 $\langle \boldsymbol{x}^*, \boldsymbol{v} \rangle > \beta$. 我们证明

$$\langle \boldsymbol{x}^*, \boldsymbol{x}^2 \rangle \geqslant \beta \quad \forall \boldsymbol{x}^2 \in S_2. \tag{1.3.10}$$

事实上,反之,若假设存在某 $\bar{\boldsymbol{x}}^2 \in S_2$,使 $\langle \boldsymbol{x}^*, \bar{\boldsymbol{x}}^2 \rangle < \beta$,则因 $\boldsymbol{x}^*$ 是线性泛函,故存在 $\boldsymbol{z} \in (\boldsymbol{v}, \bar{\boldsymbol{x}}^2)$,使 $\langle \boldsymbol{x}^*, \boldsymbol{z} \rangle = \beta$,从而 $\boldsymbol{z} \in H$. 另外,从 $\boldsymbol{v} \in S_2^i$ 和 $\bar{\boldsymbol{x}}^2 \in S_2 \subset S_2^a$ 由定理 1.2.1 的(1)有 $[\boldsymbol{v}, \bar{\boldsymbol{x}}^2) \subset S_2^i$,故 $\boldsymbol{z} \in S_2^i$. 于是,再由 $\boldsymbol{z} \in H$ 得知存在 $\boldsymbol{z} \in H \cap S_2^i$,这导致与(1.3.9)矛盾. 以下再证明有

$$\langle \boldsymbol{x}^*, \boldsymbol{x}^1 \rangle \leqslant \beta \quad \forall \boldsymbol{x}^1 \in S_1. \tag{1.3.11}$$

仍用反证法,假设存在 $\bar{\boldsymbol{x}}^1 \in S_1$ 使 $\langle \boldsymbol{x}^*, \bar{\boldsymbol{x}}^1 \rangle > \beta$. 从 $\bar{\boldsymbol{x}}^1 \in S_1 \subset T_1$ 和 $\boldsymbol{v} \in S_2^i \subset T_2^i$,由定理 1.2.3 的(4)可知存在 $\boldsymbol{z}' \in [\bar{\boldsymbol{x}}^1, \boldsymbol{v}) \cap H$,依据(1.3.8)即有 $\boldsymbol{z}' \in [\bar{\boldsymbol{x}}^1, \boldsymbol{v})$ 使 $\langle \boldsymbol{x}^*, \boldsymbol{z}' \rangle = \beta$. 但已知 $\langle \boldsymbol{x}^*, \bar{\boldsymbol{x}}^1 \rangle > \beta$ 和 $\langle \boldsymbol{x}^*, \boldsymbol{v} \rangle > \beta$,并且 $\boldsymbol{z}' \in [\bar{\boldsymbol{x}}^1, \boldsymbol{v})$,因此 $\langle \boldsymbol{x}^*, \boldsymbol{z}' \rangle > \beta$,而这与 $\langle \boldsymbol{x}^*, \boldsymbol{z}' \rangle = \beta$ 相矛盾. 由(1.3.10)和(1.3.11)得(1.3.2),故

按定义 1.3.2 的(1)知 H 分离 S_1 和 S_2. 又由(1.3.9)知 $S_2 \not\subset H$, 所以 H 正常分离 S_1 和 S_2.

(2) 从(1.3.9)和(1.3.8)可知

$$\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle \neq \beta \quad \forall \boldsymbol{x} \in S_2^i.$$

再由(1.3.10)我们有

$$\langle \boldsymbol{x}^*, \boldsymbol{x}^2 \rangle > \beta \quad \forall \boldsymbol{x}^2 \in S_2^i.$$

于是，注意到(1.3.11)，由定义 1.3.2 的(2)即知 H 严格分离 S_1 和 S_2^i. □

定理 1.3.5 设 $S_1, S_2 \subset \mathscr{Y}$ 是凸集，$S_1^{ri} \neq \varnothing$, $S_2^{ri} \neq \varnothing$, 并且 $S_1^{ri} \cap S_2^{ri} = \varnothing$.

(1) 存在超平面 $H \subset \mathscr{Y}$ 分离 S_1 和 S_2.

(2) 存在超平面 $H \subset \mathscr{Y}$ 严格分离 S_1 和 S_2^{ri} 或 S_1^{ri} 和 S_2.

证明 令 $S = S_2 - S_1$, 因为 S_1 和 S_2 是凸的，由定理 1.1.2 中的(1)知 S 是凸的. 据此，由定义 1.2.1 的(2)和定理 1.2.1 的(2)，可以推知

$$S^{ri} = (S_2 - S_1)^{ri} = S_2^{ri} - S_1^{ri}.$$

于是，由 $S_1^{ri} \cap S_2^{ri} = \varnothing$ 得

$$\boldsymbol{0} \notin S^{ri}. \tag{1.3.12}$$

(1) 设 $\boldsymbol{0} \notin \mathrm{aff} S$. 由定义 1.1.6 和定义 1.1.3 可知，存在 $\bar{\boldsymbol{z}} \in S$ 和 $\boldsymbol{z}_0 \notin \mathrm{lin}(S - \bar{\boldsymbol{z}})$ 使

$$\mathrm{aff} S = \boldsymbol{z}_0 + \mathrm{lin}(S - \bar{\boldsymbol{z}}). \tag{1.3.13}$$

因为 $\mathrm{lin}(S - \bar{\boldsymbol{z}}) \subset \mathscr{Y}$ 是维数不为 0 的真线性子空间，设 $\{\boldsymbol{e}_i\}_{i \in I}$ 是 $\mathrm{lin}(S - \bar{\boldsymbol{z}})$ 的 Hamel 基[18]，则 $\{\boldsymbol{z}_0\} \cup \{\boldsymbol{e}_i\}_{i \in I}$ 可扩充为 $\mathscr{Y}$ 的基，记作 $\{\boldsymbol{z}_0\} \cup \{\boldsymbol{e}_j\}_{j \in J} (I \subset J)$. 由于任何 $\boldsymbol{z}^* \in \mathscr{Y}^* \setminus \{\boldsymbol{0}\}$ 都可由实数组 $\{\langle \boldsymbol{z}^*, \boldsymbol{z}_0 \rangle\} \cup \{\langle \boldsymbol{z}^*, \boldsymbol{e}_j \rangle\}_{j \in J}$ 唯一确定，所以存在 $\boldsymbol{x}^* \in \mathscr{Y}^* \setminus \{\boldsymbol{0}\}$, 使

$$\langle x^*, z_0 \rangle = 1 \tag{1.3.14}$$

和

$$\langle x^*, e_i \rangle = 0,\ i \in I \subset J. \tag{1.3.15}$$

又因为任意的 $z \in \mathrm{lin}(S - \bar{z})$ 是 $\{e_i\}_{i\in I}$ 的线性组合，故由(1.3.15)得

$$\langle x^*, z \rangle = 0 \quad \forall z \in \mathrm{lin}(S - \bar{z}). \tag{1.3.16}$$

从(1.3.13)可知，对任意的 $x \in S \subset \mathrm{aff}S$，存在 $z' \in \mathrm{lin}(S - \bar{z})$ 使 $x = z_0 + z'$，故由(1.3.14)和(1.3.16)知

$$\langle x^*, x \rangle = \langle x^*, z_0 \rangle + \langle x^*, z' \rangle = 1.$$

因为 $S = S_2 - S_1$，令 $x = x^2 - x^1$，$x^1 \in S_1$，$x^2 \in S_2$，由 $x \in S$ 的任意性，从上式有 $\langle x^*, x^2 - x^1 \rangle = 1$，或

$$\langle x^*, x^1 \rangle = \langle x^*, x^2 \rangle - 1 \quad \forall x^1 \in S_1, \forall x^2 \in S_2.$$

令 $\beta = \sup\limits_{x^1 \in S_1} \langle x^*, x^1 \rangle + \dfrac{1}{2}$，则得到

$$\langle x^*, x^1 \rangle < \beta < \langle x^*, x^2 \rangle \quad \forall x^1 \in S_1, \forall x^2 \in S_2. \tag{1.3.17}$$

由此，根据定义 1.3.2 的(1)，结论成立.

设 $\mathbf{0} \in \mathrm{aff}S$. 这时 $\mathscr{X}_0 = \mathrm{aff}S = \mathrm{lin}S$ 是 $\mathscr{X}$ 的线性子空间，并且由定义 1.2.1 的(1)和(2)知 S^{ri} 即 S 在 $\mathscr{X}_0$ 上的代数内部. 因此，由(1.3.12)和注 1.2.1，利用定理 1.3.4 的(1)得知，存在超平面分离 $\{\mathbf{0}\}$ 和 S，即存在 $x_1^* \in \mathscr{X}_0^* \setminus \{\mathbf{0}\}$ 和 $\beta_1 \in R$ 使得

$$0 \leqslant \beta_1 \leqslant \langle x_1^*, y \rangle \quad \forall y \in S.$$

将 x_1^* 延拓到整个空间，则存在 $x^* \in \mathscr{X}^* \setminus \{\mathbf{0}\}$ 有

$$0 \leqslant \langle x_1^*, y \rangle = \langle x^*, y \rangle \quad \forall y \in S.$$

令 $y = x^2 - x^1$，$x^1 \in S_1$，$x^2 \in S_2$，从上式得到

$$\langle x^*, x^1\rangle \leqslant \langle x^*, x^2\rangle \quad \forall x^1 \in S_1, \forall x^2 \in S_2.$$

取 $\beta = \sup\limits_{x^1\in S_1}\langle x^*, x^1\rangle$，由定义 1.3.2 的(1)即得知 H 分离 S_1 和 S_2.

(2) 设 $\boldsymbol{0} \notin \mathrm{aff}S$. 由(1.3.17)和 $S_1^{ri} \subset S_1$，$S_2^{ri} \subset S_2$，按定义 1.3.2 中的(2)知结论成立.

设 $\boldsymbol{0} \in \mathrm{aff}S$. 与(1)中的证明相同，由(1.3.12)利用定理 1.3.4 中的(2)得知，存在 $x_2^* \in \mathscr{Y}_0^*\setminus\{\boldsymbol{0}\}$ 和 $\beta_2 \in R$ 使得

$$0 \leqslant \beta_2 < \langle x_2^*, y\rangle \quad \forall y \in S^{ri}.$$

将 x_2^* 延拓到整个空间，则存在 $x^* \in \mathscr{Y}^*\setminus\{\boldsymbol{0}\}$ 有

$$0 \leqslant \beta_2 < \langle x_2^*, y\rangle = \langle x^*, y\rangle \quad \forall y \in S^{ri}.$$

令 $y = x^2 - x^1$，$x^1 \in S_1^{ri}$，$x^2 \in S_2^{ri}$，从 $S^{ri} = S_2^{ri} - S_1^{ri}$ 和上式，有

$$\langle x^*, x^1\rangle < \langle x^*, x^2\rangle \quad \forall x^1 \in S_1^{ri}, \forall x^2 \in S_2^{ri}. \qquad (1.3.18)$$

取 $\beta' = \sup\limits_{x^1\in S_1}\langle x^*, x^1\rangle$，$\beta'' = \inf\limits_{x^2\in S_2}\langle x^*, x^2\rangle$，则有

$$\langle x^*, x^1\rangle \leqslant \beta' \leqslant \beta'' \leqslant \langle x^*, x^2\rangle \quad \forall x^1 \in S_1^{ri}, \forall x^2 \in S_2.$$

若 $\beta' < \beta''$，则

$$\langle x^*, x^1\rangle \leqslant \beta' < \langle x^*, x^2\rangle \quad \forall x^1 \in S_1^{ri}, \forall x^2 \in S_2. \qquad (1.3.19)$$

若 $\beta' = \beta''$，则不存在 $\bar{x}^1 \in S_1^{ri}$ 和 $\bar{x}^2 \in S_2^{ri}$ 使

$$\langle x^*, \bar{x}^1\rangle = \beta' = \beta'' = \langle x^*, \bar{x}^2\rangle,$$

否则，与(1.3.18)相矛盾. 因此得知或者有(1.3.19)，或者有

$$\langle x^*, x^1\rangle < \beta'' \leqslant \langle x^*, x^2\rangle \quad \forall x^1 \in S_1, \forall x^2 \in S_2^{ri}.$$

于是，由定义 1.3.2 的(2)即得 H 严格分离 S_1 和 S_2^{ri} 或 S_1^{ri} 和 S_2. □

注 1.3.2 因为 $S_1^{ri}\subset S_1$，$S_2^{ri}\subset S_2$，所以在定理 1.3.5 的条件下，由定理 1.3.5 的(2)还可推知存在超平面 $H\subset\mathscr{Y}$ 严格分离 S_1^{ri}

和 S_2^{ri}.

定理 1.3.6 设集合 $S_1, S_2 \subset \mathscr{Y}$ 非空. 若 $S_1 - S_2$ 是凸集, $(S_1 - S_2)^{ri} \neq \varnothing$, 并且 $\boldsymbol{0} \notin (S_1 - S_2)^a$, 则存在超平面 $H \subset \mathscr{Y}$ 强分离 S_1 和 S_2.

证明 令 $S = S_1 - S_2$, 则 $S^{ri} \neq \varnothing$, $\boldsymbol{0} \notin S^a$.

设 $\boldsymbol{0} \notin \mathrm{aff} S$. 由定理 1.3.5 的证明可知,存在 $\boldsymbol{x}^* \in \mathscr{Y}^* \setminus \{\boldsymbol{0}\}$ 使

$$\langle \boldsymbol{x}^*, \boldsymbol{x}^1 \rangle = \langle \boldsymbol{x}^*, \boldsymbol{x}^2 \rangle - 1 \quad \forall \boldsymbol{x}^1 \in S_1, \forall \boldsymbol{x}^2 \in S_2.$$

由此,按定义 1.3.2 的(3)可以推知结论成立.

设 $\boldsymbol{0} \in \mathrm{aff} S$. 取 $\boldsymbol{v} \in S^{ri}$, 因为 $\boldsymbol{0} \notin S^a$, 则存在 $\bar{\boldsymbol{v}} \in (\boldsymbol{0}, \boldsymbol{v})$ 使 $[\bar{\boldsymbol{v}}, \boldsymbol{v}) \subset S^a$, $[\boldsymbol{0}, \bar{\boldsymbol{v}}) \notin S^a$, 故有 $\bar{\boldsymbol{v}} \notin S^{ri}$. 据此,视$\{\bar{\boldsymbol{v}}\}$为单点凸集,又因 S 是凸的,由定理 1.2.2 的(2)知 S^{ri}是凸的,于是由定理 1.3.5 的(2)知存在 $\boldsymbol{x}^* \in \mathscr{Y}^* \setminus \{\boldsymbol{0}\}$ 使

$$\langle \boldsymbol{x}^*, \boldsymbol{v} \rangle < \langle \boldsymbol{x}^*, \bar{\boldsymbol{v}} \rangle \quad \forall \boldsymbol{v} \in S^{ri},$$

从而得

$$\sup_{\boldsymbol{v} \in S} \langle \boldsymbol{x}^*, \boldsymbol{v} \rangle \leqslant \langle \boldsymbol{x}^*, \bar{\boldsymbol{v}} \rangle. \tag{1.3.20}$$

由于 $\bar{\boldsymbol{v}} \in (\boldsymbol{0}, \boldsymbol{v})$, 则存在 $\lambda \in (0, 1)$ 使 $\bar{\boldsymbol{v}} = \lambda \boldsymbol{v}$, 又由 $\boldsymbol{v} \in S^{ri}$, 故

$$\langle \boldsymbol{x}^*, \boldsymbol{v} \rangle < \langle \boldsymbol{x}^*, \bar{\boldsymbol{v}} \rangle = \lambda \langle \boldsymbol{x}^*, \boldsymbol{v} \rangle,$$

从而有 $\langle \boldsymbol{x}^*, \boldsymbol{v} \rangle < 0$ 和 $\langle \boldsymbol{x}^*, \bar{\boldsymbol{v}} \rangle < 0$. 根据(1.3.20),则得 $\sup\limits_{\boldsymbol{v} \in S} \langle \boldsymbol{x}^*, \boldsymbol{v} \rangle < 0$. 由 $S = S_1 - S_2$, 令 $\boldsymbol{v} = \boldsymbol{x}^1 - \boldsymbol{x}^2$, $\boldsymbol{x}^1 \in S_1$, $\boldsymbol{x}^2 \in S_2$, 我们有

$$\sup_{\substack{\boldsymbol{x}^1 \in S_1 \\ \boldsymbol{x}^2 \in S_2}} \langle \boldsymbol{x}^*, \boldsymbol{x}^1 - \boldsymbol{x}^2 \rangle < 0,$$

因而得到

$$\sup_{\boldsymbol{x}^1 \in S_1} \langle \boldsymbol{x}^*, \boldsymbol{x}^1 \rangle < \inf_{\boldsymbol{x}^2 \in S_2} \langle \boldsymbol{x}^*, \boldsymbol{x}^2 \rangle.$$

据此,由定义 1.3.2 中的(3)得知结论成立. ▯

以下介绍线性拓扑空间 $\mathscr{X}$ 中的凸集分离定理.

设 $\mathscr{X}^*$ 是 $\mathscr{X}$ 的对偶空间，即 $\mathscr{X}^*$ 是 $\mathscr{X}$ 上连续线性泛函全体构成的空间. 记 $\langle x^*, x\rangle$ 表示连续线性泛函 $x^* \in \mathscr{X}^*$ 在点 $x \in \mathscr{X}$ 处的值. 可以证明，当 $x^* \in \mathscr{X}^*\backslash\{\boldsymbol{0}\}$ 和 $\beta\in R$ 时，$\mathscr{X}$ 中由连续线性泛函确定的超平面

$$H = \{x \in \mathscr{X} \mid \langle x^*, x\rangle = \beta\}$$

是闭集[14]. 这时，H 是 $\mathscr{X}$ 中的闭超平面.

定理 1.3.7 设 $S_1, S_2\subset\mathscr{X}$ 是凸集，$S_1 \neq \varnothing$，$\mathrm{int}S_2 \neq \varnothing$，并且 $S_1 \cap \mathrm{int}S_2 = \varnothing$.

(1) 存在闭超平面 $H\subset\mathscr{X}$ 正常分离 S_1 和 S_2.

(2) 存在闭超平面 $H\subset\mathscr{X}$ 严格分离 S_1 和 $\mathrm{int}S_2$.

证明 由 $\mathrm{int}S_2 \neq \varnothing$ 和定理 1.2.8 的(1)有 $\mathrm{int}S_2 = S_2^i$. 注意到由 $\mathscr{X}$ 上的连续线性泛函确定的超平面是闭的，利用定理 1.3.4 即可得到结论. ▯

定理 1.3.8 设 $\mathscr{X}$ 是局部凸线性拓扑空间，$S_1, S_2\subset\mathscr{X}$ 是凸集，$\mathrm{ri}S_1 \neq \varnothing$，$\mathrm{ri}S_2 \neq \varnothing$，并且 $\mathrm{ri}S_1 \cap \mathrm{ri}S_2 = \varnothing$.

(1) 存在闭超平面 $H\subset\mathscr{X}$ 分离 S_1 和 S_2.

(2) 存在闭超平面 $H\subset\mathscr{X}$ 严格分离 $\mathrm{ri}S_1$ 和 $\mathrm{ri}S_2$.

证明 由 $\mathrm{ri}S_1\neq\varnothing$ 和 $\mathrm{ri}S_2\neq\varnothing$，根据定理 1.2.8 的(2)，我们有 $\mathrm{ri}S_1 = S_1^{ri}$ 和 $\mathrm{ri}S_2 = S_2^{ri}$. 于是，由定理 1.3.5 和注 1.3.2，注意到这时超平面是闭的即得结论. ▯

定理 1.3.9 设 $\mathscr{X}$ 是分离局部凸线性拓扑空间，$S_1\subset\mathscr{X}$ 是紧凸集，$S_2\subset\mathscr{X}$ 是闭凸集. 若 $S_1 \cap S_2 = \varnothing$，则存在闭超平面强分离 S_1 和 S_2.

证明 因为 $S_1 - S_2$ 是闭凸集，而 $\boldsymbol{0}\notin S_1 - S_2$，从而存在凸零邻域 U，使得 $U \cap (S_1 - S_2) = \varnothing$. 由此，根据文献[25]的定理 3.2.10 即可得证. ▯

定理 1.3.10 设 $S_1, S_2\subset\mathscr{X}$ 是开凸集，$S_1 \neq \varnothing$，$S_2 \neq \varnothing$. 若 $S_1 \cap S_2 = \varnothing$，则存在闭超平面严格分离 S_1 和 S_2.

证明 因 S_2 是开凸集，故 $\mathrm{int}S_2 = S_2$. 由 $S_1 \neq \varnothing$，$\mathrm{int}S_2 = S_2$

$\neq \varnothing$ 和 $S_1 \cap \mathrm{int}S_2 = S_1 \cap S_2 = \varnothing$，根据定理 1.3.7 的(2)可得结论. □

最后，利用凸集分离定理，阐述凸集在一点处被超平面支撑以及闭凸集由它的支撑函数表示的问题.

定义 1.3.3 设集合 $S \subset \mathscr{Y}$ 非空，点 $\boldsymbol{x}^0 \in S^b$，$H \subset \mathscr{Y}$ 是超平面. 若 $\boldsymbol{x}^0 \in H$，并且 $S \subset H_+^c$ 或 H_-^c，则称 H 是集合 S 在点 $\boldsymbol{x}^0$ 处的支撑超平面，或称 H 在点 $\boldsymbol{x}^0$ 处支撑 S，$\boldsymbol{x}^0$ 称为支撑点. 再若有 $S \not\subset H$，则称 H 在点 $\boldsymbol{x}^0$ 处正常(非平凡)支撑 S；若有 $S \subset H$，则称 H 在点 $\boldsymbol{x}^0$ 处奇异(平凡)支撑 S.

图 1.3.2 的(a)和(b)分别是正常支撑和奇异支撑的情况.

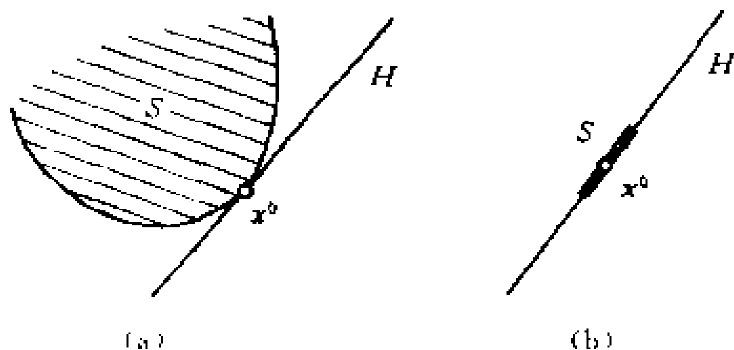


图 1.3.2

推论 1.3.11 设 $S \subset \mathscr{Y}$ 是凸集，$S^{ri} \neq \varnothing$，则对任意的 $\boldsymbol{x}^0 \in S \backslash S^{ri}$，存在超平面 $H \subset \mathscr{Y}$，它在点 $\boldsymbol{x}^0$ 处正常支撑 S.

证明 设 $S_1 = \{\boldsymbol{x}^0\}$，$S_2 = S$，由定理的条件有 $S_1^{ri} \cap S_2^{ri} \subset \{\boldsymbol{x}^0\} \cap S^{ri} = \varnothing$. 于是，利用定理 1.3.5 的(1)得知存在超平面 $H \subset \mathscr{Y}$ 分离 $\{\boldsymbol{x}^0\}$ 和 S. 注意到 $\{\boldsymbol{x}^0\}$ 是单点集，则有 $\boldsymbol{x}^0 \in H$，又 $S^{ri} \neq \varnothing$，故 $S \not\subset H$. 因此，由定义 1.3.3 即得结论. □

推论 1.3.12 设 $S \subset \mathscr{X}$ 是凸集，$\mathrm{int}S \neq \varnothing$，则对任意的点 $\boldsymbol{x}^0 \in \mathrm{rb}S$，存在闭超平面 $H \subset \mathscr{X}$，它在点 $\boldsymbol{x}^0$ 处正常支撑 S.

证明 与推论 1.3.11 的证明类似，利用定理 1.3.7 的(1)可得证. □

定义 1.3.4 设集合 $S \subset \mathscr{Y}$ 非空，并且 $S^{ri} \neq \varnothing$. 记

$$\sigma_S(x^*) = \sup_{x\in S}\langle x^*, x\rangle,\ x^*\in\mathscr{X}^*,$$

则称 $\sigma_S: \mathscr{X}^* \to R\cup\{+\infty\}$是集合 S 的支撑函数.

定理 1.3.13 设集合 $S\subset\mathscr{X}$ 非空，$\mathrm{ri}S\neq\varnothing$，$\sigma_S: \mathscr{X}^*\to R\cup\{+\infty\}$ 是集合 S 的支撑函数，则

$$\overline{\mathrm{co}}S = \{x\in\mathscr{X} \mid \langle x^*, x\rangle \leqslant \sigma_S(x^*)\quad \forall x^*\in\mathscr{X}^*\}. \tag{1.3.21}$$

证明 记(1.3.21)右端的集合为 D，先证 $\overline{\mathrm{co}}S\subset D$. 不难验证 D 是闭集，故据定理 1.2.10，由于 $\overline{\mathrm{co}}S = \mathrm{cl}(\mathrm{co}S)$，只需证明 $\mathrm{co}S\subset D$ 即可. 为此，任取 $x\in\mathrm{co}S$，由定理 1.1.8 知存在正整数 m，有 $x = \sum_{i=1}^{m}\lambda_i x^i$，其中 $x^i\in S$，$\lambda_i\geqslant 0$ $(i=1,\cdots,m)$，$\sum_{i=1}^{m}\lambda_i = 1$. 因 σ_S 是 S 的支撑函数，由定义 1.3.4 有

$$\langle x^*, x^i\rangle \leqslant \sup_{x^i\in S}\langle x^*, x^i\rangle = \sigma_S(x^*),\ x^*\in\mathscr{X}^*.$$

据此，得知

$$\langle x^*, x\rangle = \left\langle x^*, \sum_{i=1}^{m}\lambda_i x^i\right\rangle = \sum_{i=1}^{m}\lambda_i\langle x^*, x^i\rangle$$

$$\leqslant \sum_{i=1}^{m}\lambda_i\sigma_S(x^*) = \sigma_S(x^*)\sum_{i=1}^{m}\lambda_i = \sigma_S(x^*),$$

从而 $x\in D$.

以下证 $D\subset\overline{\mathrm{co}}S$. 用反证法，假设存在 $\bar{x}\in D$，但 $\bar{x}\notin\overline{\mathrm{co}}S$，则从前者有

$$\langle x^*, \bar{x}\rangle \leqslant \sigma_S(x^*)\quad \forall x^*\in\mathscr{X}^*, \tag{1.3.22}$$

而从后者(由于 $\mathrm{ri}S\neq\varnothing$) 得 $\mathrm{ri}(\overline{\mathrm{co}}S - \{\bar{x}\})\neq\varnothing$ 和 $\mathbf{0}\notin\mathrm{cl}(\overline{\mathrm{co}}S - \{\bar{x}\})$. 因此，由定理 1.3.9 的(1)得知，存在闭超平面强分离$\overline{\mathrm{co}}S$ 和 $\{\bar{x}\}$，按定义 1.3.2 的(3)也即存在 $\bar{x}^*\in\mathscr{X}^*\setminus\{\mathbf{0}\}$ 使

$$\sup_{x\in\overline{\mathrm{co}}S}\langle x^*, x\rangle < \langle\bar{x}^*, \bar{x}\rangle.$$

由定义 1.3.4 和上式得

$$\sigma_S(\overline{x}^*) = \sup_{x\in S}\langle \overline{x}^*, x\rangle \leqslant \sup_{x\in \overline{\mathrm{co}}S}\langle \overline{x}^*, x\rangle < \langle \overline{x}^*, \overline{x}\rangle,$$

这导致与(1.3.22)矛盾. □

注 1.3.3　若 $S\subset\mathscr{X}$ 是闭凸集, $\mathrm{int}S\neq\varnothing$, 由定理 1.3.13 有

$$S = \{x\in\mathscr{X} \mid \langle x^*, x\rangle \leqslant \sigma_S(x^*) \quad \forall x^*\in\mathscr{X}^*\}.$$

这表明任一内部为非空的闭凸集均可由它的支撑函数来表示.

对于 $x^*\in\mathscr{X}^*$, 考虑超平面

$$H = \{x\in\mathscr{X} \mid \langle x^*, x\rangle = \sigma_S(x^*)\}. \tag{1.3.23}$$

设有点 $x^0\in\mathscr{X}$ 使 $\langle x^*, x^0\rangle = \sigma_S(x^*)$, 则 $x^0\in H$. 从(1.3.21)和(1.3.23)可知 $S\subset\overline{\mathrm{co}}S\subset H_-^C$. 因此, 由定义 1.3.3 得知(1.3.23)就是集合 S 在点 x^0 处的支撑超平面.

§1.4　凸锥和对偶锥

凸锥和它的对偶锥是两类特殊的凸集. 它们是凸分析和非光滑分析研究中的重要概念和工具.

先介绍线性空间 $\mathscr{Y}$ 中的凸锥, 并讨论它的有关性质.

定义 1.4.1　设集合 $K\subset\mathscr{Y}$. 若对任意的 $\lambda>0$, 有

$$\lambda x\in K \quad \forall x\in K,$$

则称 K 是($\mathscr{Y}$ 中的)锥. 若 $\boldsymbol{0}\in K$, 则称 K 是有锋锥, 否则, 称 K 是无锋锥.

我们约定: 空集是锥.

定义 1.4.2　设 $K\subset\mathscr{Y}$ 是锥.

(1) 若对任意的 $x\in K\setminus\{\boldsymbol{0}\}$, 有 $-x\notin K$, 则称 K 是尖锥.

(2) 若 K 是闭集, 则称 K 是闭锥.

(3) 若 K 是凸集, 则称 K 是凸锥.

例 1.4.1 以下集合是相应空间中的凸锥.

Euclid 空间 R^n 中的非负锥:

$$R^n_+ = \{x \in R^n \mid x \geqslant \boldsymbol{0}\}$$

是 R^n 中的有锋尖闭凸锥;正锥:

$$\mathrm{int}R^n_+ = \{x \in R^n \mid x > \boldsymbol{0}\}$$

是 R^n 中的无锋尖凸锥.

Banach 空间 $\mathscr{B}$ 中过原点的超平面:

$$H(x^*, \boldsymbol{0}) = \{x \in \mathscr{B} \mid \langle x^*, x\rangle = 0, x^* \in \mathscr{B}^* \setminus \{\boldsymbol{0}\}\}$$

是 $\mathscr{B}$ 中的有锋闭凸锥,其中 $\mathscr{B}^*$ 是 $\mathscr{B}$ 的对偶空间.

线性空间 $\mathscr{Y}$ 中过原点的开半空间的交:

$$K = \{x \in \mathscr{Y} \mid \langle x_i^*, x\rangle > 0, x_i^* \in \mathscr{Y}^* \setminus \{\boldsymbol{0}\}, i \in I\}$$

是 $\mathscr{Y}$ 中的无锋凸锥(其中 I 是任意指标集).

定理 1.4.1 设集合 $K \subset \mathscr{Y}$. K 是凸锥当且仅当对任意的 λ, $\mu > 0$, 有

$$\lambda x^1 + \mu x^2 \in K \quad \forall x^1, x^2 \in K. \tag{1.4.1}$$

证明 必要性. 设 K 是凸锥,则它是凸集. 于是,由定义 1.1.1 可知,对任意的 λ, $\mu > 0$,因为 $\frac{\lambda}{\lambda+\mu}$, $\frac{\mu}{\lambda+\mu} \in (0, 1)$, 我们有

$$\frac{\lambda}{\lambda+\mu}x^1 + \frac{\mu}{\lambda+\mu}x^2 \in K \quad \forall x^1, x^2 \in K.$$

由于 K 是锥,故根据定义 1.4.1 得

$$\lambda x^1 + \mu x^2 = (\lambda+\mu)\left(\frac{\lambda}{\lambda+\mu}x^1 + \frac{\mu}{\lambda+\mu}x^2\right) \in K \quad \forall x^1, x^2 \in K.$$

充分性. 在(1.4.1)中取 $x^1 = x^2$, 则对任意的 $\lambda + \mu > 0$, 有 $(\lambda + \mu)x^1 \in K$. 因此,由定义 1.4.1 知 K 是锥. 取 $\lambda + \mu = 1$, 由 λ、

$\mu > 0$ 我们有 $\lambda \in (0, 1)$，并且

$$\lambda \boldsymbol{x}^1 + (1 - \lambda)\boldsymbol{x}^2 \in K \quad \forall \boldsymbol{x}^1, \boldsymbol{x}^2 \in K.$$

于是，按定义 1.1.1 可知 K 是凸集，再据定义 1.4.2 的(3)使得 K 是凸锥. ▯

定理 1.4.2 设 $K \subset \mathscr{Y}$ 是凸锥，则 $K + K \subset K$.

证明 设 $\boldsymbol{x} \in K + K$，则存在 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in K$ 使 $\boldsymbol{x} = \boldsymbol{x}^1 + \boldsymbol{x}^2$. 因为 K 是凸锥，故由定理 1.4.1 即得 $\boldsymbol{x} = \boldsymbol{x}^1 + \boldsymbol{x}^2 \in K$. ▯

定理 1.4.3 设 $K_i \subset \mathscr{Y}$ $(i = 1, \cdots, m)$ 是凸锥.

(1) $\bigcap\limits_{i=1}^{m} K_i$ 是凸锥.

(2) $\sum\limits_{i=1}^{m} K_i$ 是凸锥.

证明 (1) 任取 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in \bigcap\limits_{i=1}^{m} K_i$，则 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in K_i (i = 1, \cdots, m)$. 因为 $K_i (i=1, \cdots, m)$ 是凸锥，由定理 1.4.1 知，对任意的 $\lambda, \mu > 0$，有

$$\lambda \boldsymbol{x}^1 + \mu \boldsymbol{x}^2 \in K_i, \quad i = 1, \cdots, m,$$

从而 $\lambda \boldsymbol{x}^1 + \mu \boldsymbol{x}^2 \in \bigcap\limits_{i=1}^{m} K_i$. 再据定理 1.4.1，即知 $\bigcap\limits_{i=1}^{m} K_i$ 是凸锥.

(2) 任取 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in \sum\limits_{i=1}^{m} K_i$，则存在 $\boldsymbol{y}_i^1, \boldsymbol{y}_i^2 \in K_i (i = 1, \cdots, m)$，使 $\boldsymbol{x}^1 = \sum\limits_{i=1}^{m} \boldsymbol{y}_i^1$，$\boldsymbol{x}^2 = \sum\limits_{i=1}^{m} \boldsymbol{y}_i^2$. 因为 $K_i (i=1, \cdots, m)$ 是凸锥，由定理 1.4.1 知，对任意的 $\lambda, \mu > 0$，有

$$\lambda \boldsymbol{y}_i^1 + \mu \boldsymbol{y}_i^2 \in K_i, \; i = 1, \cdots, m.$$

据此，我们推知

$$\lambda \boldsymbol{x}^1 + \mu \boldsymbol{x}^2 = \sum_{i=1}^{m} (\lambda \boldsymbol{y}_i^1 + \mu \boldsymbol{y}_i^2) \in \sum_{i=1}^{m} K_i,$$

再由定理 1.4.1 即得 $\sum\limits_{i=1}^{m} K_i$ 是凸锥. ▯

定理 1.4.4 设 $\mathscr{W}$ 是线性空间，$K_1\subset\mathscr{V}$ 和 $K_2\subset\mathscr{W}$ 都是凸锥，则 $K_1\times K_2\subset\mathscr{V}\times\mathscr{W}$ 是凸锥.

证明 由定理 1.4.1 可以推证. ▯

定理 1.4.5 设 $K_1, K_2\subset\mathscr{V}$ 是非空凸锥，则

$$K_1+K_2=\mathrm{co}(K_1\cup K_2).$$

证明 先证 $K_1+K_2\subset\mathrm{co}(K_1\cup K_2)$. 任取 $\boldsymbol{x}\in K_1+K_2$，则存在 $\boldsymbol{x}^1\in K_1$，$\boldsymbol{x}^2\in K_2$，使 $\boldsymbol{x}=\boldsymbol{x}^1+\boldsymbol{x}^2$. 因为 K_1 和 K_2 是锥，故有 $2\boldsymbol{x}^1\in K_1$ 和 $2\boldsymbol{x}^2\in K_2$. 从而 $2\boldsymbol{x}^1$，$2\boldsymbol{x}^2\in K_1\cup K_2\subset\mathrm{co}(K_1\cup K_2)$. 由于 $\mathrm{co}(K_1\cup K_2)$ 是凸集，因此得

$$\boldsymbol{x}=\frac{1}{2}(2\boldsymbol{x}^1)+\frac{1}{2}(2\boldsymbol{x}^2)\in\mathrm{co}(K_1\cup K_2).$$

现证 $\mathrm{co}(K_1\cup K_2)\subset K_1+K_2$. 任取 $\boldsymbol{x}\in\mathrm{co}(K_1\cup K_2)$，因为 K_1 和 K_2 是凸集，由定理 1.1.9 的(3)知，存在 $\bar{\lambda}\in[0,1]$，有 $\boldsymbol{x}\in\bar{\lambda}K_1+(1-\bar{\lambda})K_2$. 由此，有 $\boldsymbol{x}^1\in K_1$，$\boldsymbol{x}^2\in K_2$，使 $\boldsymbol{x}=\bar{\lambda}\boldsymbol{x}^1+(1-\bar{\lambda})\boldsymbol{x}^2$. 由于 K_1 和 K_2 是锥，故 $\bar{\lambda}\boldsymbol{x}^1\in K_1$，$(1-\bar{\lambda})\boldsymbol{x}^2\in K_2$，于是得到 $\boldsymbol{x}\in K_1+K_2$. ▯

定义 1.4.3 设集合 $S\subset\mathscr{V}$，则 $\mathscr{V}$ 中所有包含 S 的有锋凸锥的交称为 S 的锥包(或由 S 生成的凸锥)，记作 $\mathrm{cone}S$.

我们约定：空集的锥包是空集.

注 1.4.1 由定义 1.4.3，集合的锥包是包含它的最小凸锥加上原点. 显然，若 $K\subset\mathscr{V}$ 是凸锥，则 $\mathrm{cone}K=K\cup\{\boldsymbol{0}\}$.

定理 1.4.6 设集合 $S\subset\mathscr{V}$，则对任何正整数 m 有

$$\mathrm{cone}S=\left\{\boldsymbol{x}\in\mathscr{V}\;\middle|\;\boldsymbol{x}=\sum_{i=1}^{m}\lambda_i\boldsymbol{x}^i\quad\forall\boldsymbol{x}^i\in S,\ \forall\lambda_i\geqslant0,\ i=1,\cdots,m\right\},$$

即 $\mathrm{cone}S$ 是 S 中向量的有限非负线性组合构成的集合.

证明 由定义 1.4.3，利用定理 1.4.1 并注意 $\mathrm{cone}S$ 是包含原点的，即可推得. ▯

定理 1.4.7 设集合 $S_1, S_2\subset\mathscr{V}$.

(1) $\mathrm{cone}(S_1 \cap S_2) \subset \mathrm{cone}S_1 \cap \mathrm{cone}S_2$. 若 S_1 和 S_2 是凸集并且 $\mathbf{0} \in S_1 \cap S_2$, 则 $\mathrm{cone}(S_1 \cap S_2) = \mathrm{cone}S_1 \cap \mathrm{cone}S_2$.

(2) $\mathrm{cone}(S_1 + S_2) \subset \mathrm{cone}S_1 + \mathrm{cone}S_2$. 若 $\mathbf{0} \in S_1 \cap S_2$, 则 $\mathrm{cone}(S_1 + S_2) = \mathrm{cone}S_1 + \mathrm{cone}S_2$.

证明 由定理 1.4.1 可以直接推证. □

利用凸锥,我们给出广义的锥凸集和严格锥凸集的概念.

定义 1.4.4 设集合 $S \subset \mathscr{Y}$ 非空, $K \subset \mathscr{Y}$ 是有锋凸锥. 若对任意的 $\lambda \in (0, 1)$, 有

$$\lambda \boldsymbol{x}^1 + (1 - \lambda)\boldsymbol{x}^2 \in S + K \quad \forall \boldsymbol{x}^1, \boldsymbol{x}^2 \in S,$$

则称 S 是($\mathscr{Y}$ 中的)K-凸集,或集合 S 是 K-凸的.

设 $\mathscr{X}$ 是线性拓扑空间.

定义 1.4.5 设集合 $S \subset \mathscr{X}$ 非空, $K \subset \mathscr{Y}$ 是内部非空的凸锥. 若对任意的 $\lambda \in (0, 1)$, 有

$$\lambda \boldsymbol{x}^1 + (1 - \lambda)\boldsymbol{x}^2 \in S + \mathrm{int}K \quad \forall \boldsymbol{x}^1, \boldsymbol{x}^2 \in S,$$

则称 S 是($\mathscr{X}$ 中的)严格 K-凸集,或集合 S 是严格 K-凸的.

注 1.4.2 由定义 1.4.4 和定义 1.4.5 易知,线性拓扑空间中的严格 K-凸集也是 K-凸集,但反之不然. 此外,由定义 1.1.1 和定义 1.4.4 可以得知,对于某凸锥 K 是 K-凸的集合,它不一定是凸集,而凸集关于任何的有锋凸锥 K 都是 K-凸的.

以下介绍线性拓扑空间 $\mathscr{X}$ 中的对偶锥.

设 $\mathscr{X}^*$ 是 $\mathscr{X}$ 的对偶空间, $\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle$ 是连续线性泛函 $\boldsymbol{x}^* \in \mathscr{X}^*$ 在点 $\boldsymbol{x} \in \mathscr{X}$ 处的值.

定义 1.4.6 设 $K \subset \mathscr{X}$ 是凸锥.

(1) 集合

$$K^* = \{\boldsymbol{x}^* \in \mathscr{X}^* \mid \langle \boldsymbol{x}^*, \boldsymbol{x} \rangle \geqslant 0 \quad \forall \boldsymbol{x} \in K\}$$

称为 K 的对偶锥.

(2) 集合

$$K^{**}=\{x\in\mathscr{X}\mid\langle x,x^*\rangle\geqslant 0\quad\forall x^*\in K^*\}$$

称为 K 的双对偶锥.

(3) 集合

$$K^+=\{x^*\in\mathscr{X}^*\mid\langle x^*,x\rangle>0\quad\forall x\in K\backslash\{\boldsymbol{0}\}\}$$

称为 K 的严格对偶锥;集合

$$K^-=\{x^*\in\mathscr{X}^*\mid\langle x^*,x\rangle\leqslant 0\quad\forall x\in K\}$$

称为 K 的负对偶锥.

注 1.4.3 由定义 1.4.6 的(1)和(2),易知 $(K^*)^*=K^{**}$. 因为 $(R^n)^*=R^n$,特别地有 $(R^n_+)^*=R^n_+$ 和 $(R^n_+)^{**}=R^n_+$. 从定义 1.4.6 中的(1)和(3),则知

$$K^+\subset K^*.$$

设 $K_1, K_2\subset\mathscr{X}$ 是凸锥. 若 $K_1\subset K_2$,由定义 1.4.6 的(1)和(3)还不难推知有 $K_2^*\subset K_1^*$ 和 $K_2^-\subset K_1^-$.

定理 1.4.8 设 $K\subset\mathscr{X}$ 是凸锥,则 K^* 和 K^{**} 是有锋闭凸锥.

证明 先证明 K^* 是有锋闭凸锥. 首先,由于对任意的 $x\in K$ 有 $\langle\boldsymbol{0},x\rangle=0$,按定义 1.4.6 的(1)知 $\boldsymbol{0}\in K^*$,故由定义 1.4.1 得 K^* 是有锋锥. 其次,设序列 $\{x_k^*\}\subset K^*$,$x_k^*\to x^*\ (k\to\infty)$,则由定义 1.4.6 的(1),对任意的 $x\in K$ 有

$$\langle x_k^*,x\rangle\geqslant 0,\ k=1,2,\cdots.$$

对上式令 $k\to\infty$,因为 $x_k^*\in\mathscr{X}^*\ (k=1,2,\cdots)$ 是 x 处的连续泛函,所以由 $x_k^*\to x^*$ 得 $\langle x^*,x\rangle\geqslant 0$. 由此,根据定义 1.4.6 的(1)知 $x^*\in K^*$,于是 K^* 是闭集. 最后,为证 K^* 是凸锥,任取 $x^1, x^2\in K^*$,由定义 1.4.6 的(1),对任意的 $x\in K$ 有 $\langle x^1,x\rangle\geqslant 0$ 和 $\langle x^2,x\rangle\geqslant 0$. 由于 $x^1, x^2\in\mathscr{X}^*$ 又是 x 处的线性泛函,故对任意的 $\lambda,\mu>0$ 有

$$\langle \lambda x^1 + \mu x^2, x \rangle = \lambda \langle x^1, x \rangle + \mu \langle x^2, x \rangle \geqslant 0.$$

据此，由定义 1.4.6 的(1)知 $\lambda x^1 + \mu x^2 \in K^*$，再由定理 1.4.1 即得 K^* 是凸锥.

与上面的证明类似，由定义 1.4.6 的(2)注意到 $x^* \in \mathscr{X}^*$ 是连续线性泛函，可得 K^{**} 是有锋闭凸锥. ▯

定理 1.4.9　设 $\mathscr{X}$ 是分离局部凸线性拓扑空间，$K \subset \mathscr{X}$ 是凸锥.

(1) $K^{**} = \mathrm{cl}K$.

(2) 若 K 是闭集，则 $K^{**} = K$.

证明　(1) 任取 $x \in K$，由定义 1.4.6 的(1)，对任意的 $x^* \in K^*$ 有 $\langle x^*, x \rangle \geqslant 0$. 于是，再由定义 1.4.6 的(2)知 $x \in K^{**}$，从而 $K \subset K^{**}$. 根据定理 1.4.8，K^{**} 是闭集，因而 $\mathrm{cl}K \subset K^{**}$. 为证 $K^{**} \subset \mathrm{cl}K$，设 $x \notin \mathrm{cl}K$，则因 $\mathrm{cl}K$ 是闭凸集，故由定理 1.3.9 知点 x 和集合 $\mathrm{cl}K$ 可强分离. 于是，按定义 1.3.2 的(3)，存在 $x^* \in \mathscr{X}^*$ 使得

$$\langle x^*, x \rangle < \inf_{y \in \mathrm{cl}K} \langle x^*, y \rangle. \tag{1.4.2}$$

显然有 $\theta \in \mathrm{cl}K$，则 $\inf\limits_{y \in \mathrm{cl}K} \langle x^*, y \rangle \leqslant 0$，因而由(1.4.2)得知

$$\langle x^*, x \rangle < 0. \tag{1.4.3}$$

设若存在 $\bar{y} \in K \setminus \{\theta\}$ 使 $\langle x^*, \bar{y} \rangle < 0$，则有 $\langle x^*, \lambda \bar{y} \rangle \to -\infty$ $(\lambda \to +\infty)$，从而得 $\inf\limits_{y \in \mathrm{cl}K} \langle x^*, y \rangle = -\infty$，与(1.4.2)相矛盾. 由此，对任意的 $y \in K \setminus \{\theta\}$ 有 $\langle x^*, y \rangle \geqslant 0$，于是

$$\langle x^*, y \rangle \geqslant 0 \quad \forall y \in K.$$

按定义 1.4.6 的(1)知 $x^* \in K^*$，再由(1.4.3)便得 $x \notin K^{**}$.

(2) 由 K 是闭集，按定义 1.2.2 的(3)有 $K = \mathrm{cl}K$. 于是，从已证得的(1)即得结论. ▯

注 1.4.4　设 K 是分离局部凸空间 $\mathscr{X}$ 中的凸锥，与定理 1.4.9 的证明类似，可以得到 $(K^-)^- = \mathrm{cl}K$. 再若 K 是闭的，则

$$(K^{*})^{*}=K.$$

定理 1.4.10 设 $K_1, K_2\subset\mathscr{X}$ 是凸锥.

(1) 若 $K_1\subset K_2$，则 $K_2^*\subset K_1^*$.

(2) $(K_1+K_2)^*=K_1^*\cap K_2^*=(K_1\cup K_2)^*$.

(3) $\mathrm{co}(K_1^*\cup K_2^*)=K_1^*+K_2^*$.

(4) $(\mathrm{cl}K_1\cap\mathrm{cl}K_2)^*=\mathrm{cl}(K_1^*+K_2^*)$.

证明 (1) 由定义 1.4.6 的(1)直接可得.

(2) 先证 $(K_1+K_2)^*=K_1^*\cap K_2^*$. 设 $\boldsymbol{x}^*\in(K_1+K_2)^*$，由定义 1.4.6 的(1)有

$$\langle\boldsymbol{x}^*,\boldsymbol{x}^1+\boldsymbol{x}^2\rangle\geqslant 0\quad\forall\boldsymbol{x}^1\in K_1,\forall\boldsymbol{x}^2\in K_2.$$

在上式中分别令 $\boldsymbol{x}^1$ 和 $\boldsymbol{x}^2$ 趋于 $\boldsymbol{0}$，则有

$$\langle\boldsymbol{x}^*,\boldsymbol{x}^1\rangle\geqslant 0\quad\forall\boldsymbol{x}^1\in K_1$$

和

$$\langle\boldsymbol{x}^*,\boldsymbol{x}^2\rangle\geqslant 0\quad\forall\boldsymbol{x}^2\in K_2.$$

由此，得到 $\boldsymbol{x}^*\in K_1^*$ 和 $\boldsymbol{x}^*\in K_2^*$，从而 $\boldsymbol{x}^*\in K_1^*\cap K_2^*$. 因上述过程可逆，故得证.

下证 $K_1^*\cap K_2^*=(K_1\cup K_2)^*$. 设 $\boldsymbol{x}^*\in(K_1\cup K_2)^*$，由定义 1.4.6 的(1)有

$$\langle\boldsymbol{x}^*,\boldsymbol{x}\rangle\geqslant 0\quad\forall\boldsymbol{x}\in K_1\cup K_2.$$

于是得知

$$\langle\boldsymbol{x}^*,\boldsymbol{x}\rangle\geqslant 0\quad\forall\boldsymbol{x}\in K_1$$

和

$$\langle\boldsymbol{x}^*,\boldsymbol{x}\rangle\geqslant 0\quad\forall\boldsymbol{x}\in K_2,$$

从而 $\boldsymbol{x}^*\in K_1^*\cap K_2^*$. 上述过程也可逆，故得证.

(3) 因为 K_1 和 K_2 是凸锥，由定理 1.4.8 知，K_1^* 和 K_2^* 是线性空间 $\mathscr{X}^*$ 中的非空凸锥. 由此，根据定理 1.4.5 即得结论.

(4) 由定理 1.4.9 的(1)，有

$$K_1^{**} = \mathrm{cl}K_1, \quad K_2^{**} = \mathrm{cl}K_2.$$

再利用已证得的(2)和定理1.4.9的(1)，可得

$$(\mathrm{cl}K_1 \cap \mathrm{cl}K_2)^* = (K_1^{**} \cap K_2^{**})^* = [(K_1^* + K_2^*)^*]^*$$
$$= (K_1^* + K_2^*)^{**} = \mathrm{cl}(K_1^* + K_2^*). \square$$

定理 1.4.11　设 $K \subset \mathscr{H}$ 是凸锥，$\mathrm{int}K \neq \varnothing$.

(1) 若 $\boldsymbol{x}^* \in K^* \backslash \{\boldsymbol{0}\}$，$\boldsymbol{x} \in \mathrm{int}K$，则 $\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle > 0$.

(2) 若 $\boldsymbol{x}^* \in \mathrm{int}K^*$，$\boldsymbol{x} \in K \backslash \{\boldsymbol{0}\}$，则 $\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle > 0$.

证明　(1) 用反证法.假设存在 $\bar{\boldsymbol{x}}^* \in K^* \backslash \{\boldsymbol{0}\}$ 和 $\bar{\boldsymbol{x}} \in \mathrm{int}K$，有

$$\langle \bar{\boldsymbol{x}}^*, \bar{\boldsymbol{x}} \rangle \leqslant 0. \tag{1.4.4}$$

由 $\mathrm{int}K$ 是开集和 $\bar{\boldsymbol{x}} \in \mathrm{int}K$，可知存在 $\bar{\boldsymbol{x}}$ 的邻域 $U(\bar{\boldsymbol{x}}) \subset K$，由定义1.4.6的(1)有

$$\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle \geqslant 0 \quad \forall \boldsymbol{x}^* \in K^*, \forall \boldsymbol{x} \in U(\bar{\boldsymbol{x}}).$$

因为 $\bar{\boldsymbol{x}}^* \in K^* \backslash \{\boldsymbol{0}\} \subset K^*$，所以从上式得

$$\langle \bar{\boldsymbol{x}}^*, \boldsymbol{x} \rangle \geqslant 0 \quad \forall \boldsymbol{x} \in U(\bar{\boldsymbol{x}}). \tag{1.4.5}$$

又由 $\bar{\boldsymbol{x}}^* \neq \boldsymbol{0}$ 知存在 $\boldsymbol{x}^0 \in \mathscr{H}$，有 $\langle \bar{\boldsymbol{x}}^*, \boldsymbol{x}^0 \rangle > 0$. 取充分小的 $\varepsilon > 0$，使 $\boldsymbol{x} = \bar{\boldsymbol{x}} - \varepsilon \boldsymbol{x}^0 \in U(\bar{\boldsymbol{x}})$，则由(1.4.5)得

$$\langle \bar{\boldsymbol{x}}^*, \bar{\boldsymbol{x}} - \varepsilon \boldsymbol{x}^0 \rangle \geqslant 0.$$

于是推知 $\langle \bar{\boldsymbol{x}}^*, \bar{\boldsymbol{x}} \rangle \geqslant \varepsilon \langle \bar{\boldsymbol{x}}^*, \boldsymbol{x}^0 \rangle > 0$，这与(1.4.4)相矛盾.

(2) 与(1)的证明类似，利用 $\mathrm{int}K^*$ 是开集即可证明. $\square$

第 2 章　凸函数和凸映射

凸函数和凸映射是凸分析的主要研究对象，其有关理论也是非光滑分析的重要基础.

本章先阐述凸函数的概念及其基本性质. 然后讨论它的连续性和对偶性问题. 在论述了两类广义意义下的凸函数之后，进而将函数的凸性概念推广到映射凸性的情况，介绍凸映射和广义凸映射.

§2.1　凸函数及有关性质

我们对取实数值的函数和实泛函统称为实值函数. 这一节介绍线性空间上具有凸性的实值函数以及它的有关性质.

设 $\mathscr{X}$ 是线性空间.

定义 2.1.1　设 $S \subset \mathscr{X}$ 是非空凸集，$f: S \to R$ 是实值函数.

(1) 若对任意的 $\lambda \in (0, 1)$ 有

$$f\left(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2\right) \leqslant \lambda f(\boldsymbol{x}^1) + (1-\lambda) f(\boldsymbol{x}^2)$$

$$\forall \boldsymbol{x}^1, \boldsymbol{x}^2 \in S, \tag{2.1.1}$$

则称 f 是集合 S 上的凸函数(凸泛函)，或函数(泛函) f 在 S 上是凸的.

(2) 若对任意的 $\lambda \in (0, 1)$ 有

$$f\left(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2\right) < \lambda f(\boldsymbol{x}^1) + (1-\lambda) f(\boldsymbol{x}^2)$$

$$\forall \boldsymbol{x}^1, \boldsymbol{x}^2 \in S,\ f(\boldsymbol{x}^1) \neq f(\boldsymbol{x}^2), \tag{2.1.2}$$

则称 f 是集合 S 上的严格凸函数，或函数 f 在 S 上是严格凸的.

(3) 若对任意的 $\lambda \in (0, 1)$ 有

$$f\left(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2\right) < \lambda f(\boldsymbol{x}^1) + (1-\lambda) f(\boldsymbol{x}^2)$$

$$\forall \boldsymbol{x}^1, \boldsymbol{x}^2 \in S, \boldsymbol{x}^1 \neq \boldsymbol{x}^2, \tag{2.1.3}$$

则称 f 是集合 S 上的强凸函数，或函数 f 在 S 上是强凸的.

若 $-f$ 是 S 上的凸函数，则称 f 是 S 上的凹函数，或 f 在 S 上是凹的. 若 $-f$ 是 S 上的严格(强)凸函数，则称 f 是 S 上的严格(强)凹函数，或 f 在 S 上是严格(强)凹的.

图 2.1.1 中的(a)、(b)、(c)和(d)依次是单变量的凸函数、严格凸函数、强凸函数和凹函数的图示.

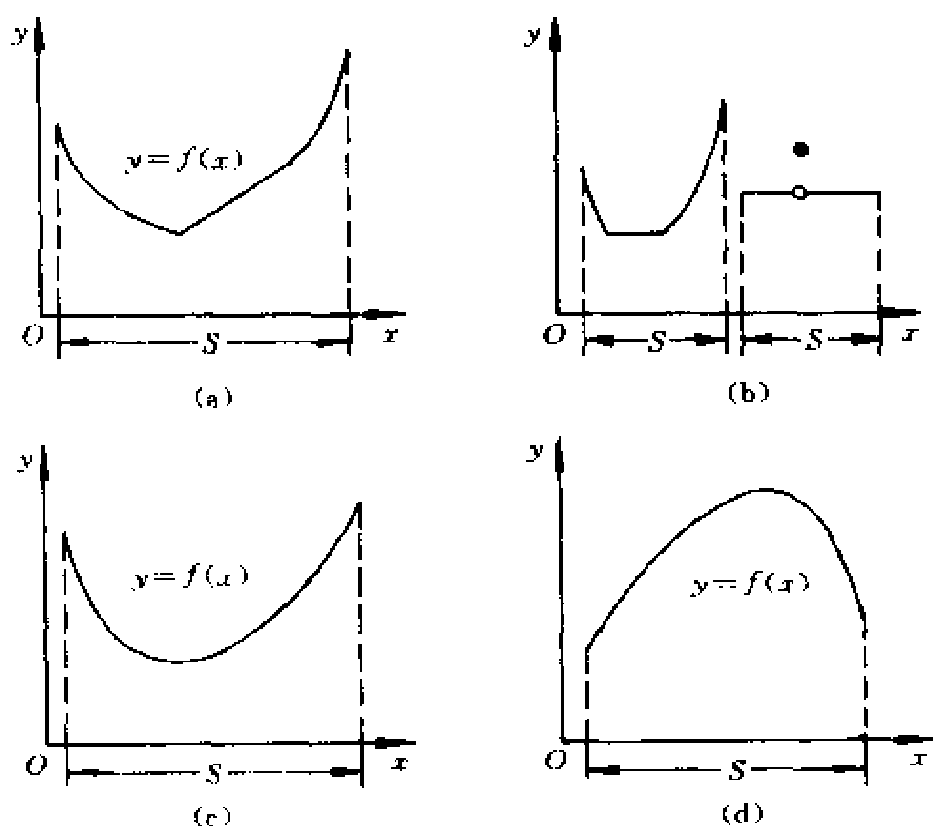


图　2.1.1

注 2.1.1　在定义 2.1.1 中固定 $\boldsymbol{x}^1 = \boldsymbol{x}^0 \in S$，令 $\boldsymbol{x}^2 = \boldsymbol{x}$，则(2.1.1)成为

$$f(\lambda x^0 + (1-\lambda)x) \leqslant \lambda f(x^0) + (1-\lambda)f(x) \quad \forall x \in S.$$

这时，称 f 在点 x^0 处相对于集合 S 是凸的. 同样地，(2.1.2)(或(2.1.3))成为

$$f(\lambda x^0 + (1-\lambda)x) < \lambda f(x^0) + (1-\lambda)f(x)$$

$$\forall x \in S, \; f(x^0) \neq f(x) (\text{或 } x^0 \neq x),$$

并称 f 在点 x^0 处相对于 S 是严格凸的(或强凸的).

注 2.1.2 从定义 2.1.1 易知，若 f 是 S 上的强凸(凹)函数，则 f 必是 S 上的凸(凹)函数和严格凸(凹)函数，但反之不然. 注意，严格凸函数不一定是凸函数(见图 2.1.1(b)).

例 2.1.1 以下函数是相应集合上的凸函数.

实数域 R 上的二次函数：

$$f(x) = x^2, \quad x \in R,$$

并且是严格凸函数和强凸函数.

Euclid 空间 R^n 上的范数函数：

$$f(x) = \|x\|_p = \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}}, \quad x \in R^n, \; p > 1,$$

其中 $x = (x_1, \cdots, x_n)^T$. 特别地，

$$f(x) = \|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$$

是 R^n 上的凸函数.

Banach 空间 $\mathscr{B}$ 中凸集 S 上的距离函数：

$$d_S(x) = \inf_{y \in S} \|x - y\|, \quad x \in \mathscr{B}.$$

线性拓扑空间 $\mathscr{X}$ 中凸集 S 上的 Minkowski 函数(泛函)：

$$\mu_S(x) = \inf\{\alpha > 0 \mid x \in \alpha S\}, \quad x \in \mathscr{X}.$$

线性空间 $\mathscr{V}$ 上的仿射函数：

$$l(x) = \langle a, x \rangle + \beta, \quad x \in \mathscr{Y},$$

其中 $a \in \mathscr{Y}$，$\beta \in R$. 它也是 $\mathscr{Y}$ 上的凹函数，而不是 $\mathscr{Y}$ 上的严格凸函数.

线性空间 $\mathscr{Y}$ 中凸集 S 上的指示函数：

$$\delta_S(x) = \begin{cases} 0, & x \in S; \\ +\infty, & x \in \mathscr{Y} \backslash S. \end{cases}$$

先给出关于凸函数的两个基本关系.

定理 2.1.1　设 $S \subset \mathscr{Y}$ 是非空凸集，则 $f: S \to R$ 是凸函数当且仅当对任何正整数 $m \geqslant 2$ 和任意的 $\lambda_i \geqslant 0$ $(i = 1, \cdots, m)$，$\sum_{i=1}^{m} \lambda_i = 1$，有

$$f\left(\sum_{i=1}^{m} \lambda_i x^i\right) \leqslant \sum_{i=1}^{m} \lambda_i f(x^i) \quad \forall x^1, \cdots, x^m \in S. \tag{2.1.4}$$

证明　充分性. 在(2.1.4)中取 $m = 2$，由定义 2.1.1 中的(1)即得.

必要性. 用数学归纳法证. 当 $m = 2$ 时，因为 f 是 S 上的凸函数，由(2.1.1)知(2.1.4)成立. 现设 $m \leqslant k$ ($k \geqslant 2$ 是正整数)时(2.1.4)成立，证明(2.1.4)对于 $m = k + 1$ 成立. 不失一般性，设 $\lambda_i > 0$ $(i = 1, \cdots, k + 1)$，$\sum_{i=1}^{k+1} \lambda_i = 1$（若有某 $\lambda_i = 0$，则 $m = k$，由假设知(2.1.4) 成立），则有 $1 - \lambda_{k+1} = \sum_{i=1}^{k} \lambda_i > 0$. 记 $\sum_{i=1}^{k} \lambda_i x^i = (1 - \lambda_{k+1}) x$，由 $x^i \in S$ $(i = 1, \cdots, k)$ 和 $\sum_{i=1}^{k} \frac{\lambda_i}{1 - \lambda_{k+1}} = 1$ 以及 S 是凸集，根据定理 1.1.4 有 $x = \sum_{i=1}^{k} \frac{\lambda_i}{1 - \lambda_{k+1}} x^i \in S$. 又因为 $x^{k+1} \in S$，以及 f 是 S 上的凸函数，按定义 2.1.1 的(1)，得知

$$f\left(\sum_{i=1}^{k+1} \lambda_i x^i\right) = f\left((1 - \lambda_{k+1}) x + \lambda_{k+1} x^{k+1}\right)$$

$$\leqslant (1-\lambda_{k+1})f(\boldsymbol{x}) + \lambda_{k+1}f(\boldsymbol{x}^{k+1})$$

$$= (1-\lambda_{k+1})f\left(\frac{1}{1-\lambda_{k+1}}\sum_{i=1}^{k}\lambda_i\boldsymbol{x}^i\right) + \lambda_{k+1}f(\boldsymbol{x}^{k+1}).$$

由于假设(2.1.4)对于 $m=k$ 成立，故有

$$f\left(\frac{1}{1-\lambda_{k-1}}\sum_{i=1}^{k}\lambda_i\boldsymbol{x}^i\right) \leqslant \sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}}f(\boldsymbol{x}^i).$$

将它代入上式，得到

$$f\left(\sum_{i=1}^{k+1}\lambda_i\boldsymbol{x}^i\right) \leqslant (1-\lambda_{k-1})\sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}}f(\boldsymbol{x}^i) + \lambda_{k+1}f(\boldsymbol{x}^{k+1})$$

$$= \sum_{i=1}^{k-1}\lambda_i f(\boldsymbol{x}^i),$$

即(2.1.4)对于 $m=k+1$ 成立. ▯

定理 2.1.2 设 $S \subset \mathscr{X}$ 是非空凸集，$f: S \to R$ 是凸函数，则对任意的 $c \in R$，水平集 $H_S(f, c) = \{\boldsymbol{x} \in S \mid f(\boldsymbol{x}) \leqslant c\}$ 是凸集.

证明 任取 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in H_S(f, c) \subset S$，因为 S 是凸集，故对任意的 $\lambda \in (0, 1)$ 有 $\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 \in S$. 由于 f 是 S 上的凸函数，按定义 2.1.1 的(1)，得

$$f(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \leqslant \lambda f(\boldsymbol{x}^1) + (1-\lambda)f(\boldsymbol{x}^2)$$

$$\leqslant \lambda c + (1-\lambda)c = c,$$

于是 $\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 \in H_S(f, c)$. 据此，由定义 1.1.1 即知 $H_S(f, c)$是凸集. ▯

为了讨论凸函数和严格凸函数之间的关系，我们引入实值函数(实泛函)的半连续性概念如下.

定义 2.1.2 设 $\mathscr{X}$ 是线性拓扑空间，$f: \mathscr{X} \to R \cup \{\pm\infty\}$ 是广义实值函数，点 $\boldsymbol{x}^0 \in \mathscr{X}$.

(1) 若对任意的 $\varepsilon > 0$，存在 $\boldsymbol{x}^0$ 的邻域 $U(\boldsymbol{x}^0)$有

$$-\varepsilon < f(\boldsymbol{x}) - f(\boldsymbol{x}^0) \quad \forall \boldsymbol{x} \in U(\boldsymbol{x}^0),$$

则称 f 在点 $\boldsymbol{x}^0$ 处是下半连续的. 若 f 在集合 $S \subset \mathscr{X}$ 的每一点处都是下半连续的，则称 f 是 S 上的下半连续函数（泛函），或 f 在 S 上是下半连续的.

(2) 若对任意的 $\varepsilon > 0$，存在 $\boldsymbol{x}^0$ 的邻域 $U(\boldsymbol{x}^0)$ 有

$$f(\boldsymbol{x}) - f(\boldsymbol{x}^0) < \varepsilon \quad \forall \boldsymbol{x} \in U(\boldsymbol{x}^0),$$

则称 f 在点 $\boldsymbol{x}^0$ 处是上半连续的. 若 f 在集合 $S \subset \mathscr{X}$ 的每一点处都是上半连续的，则称 f 是 S 上的上半连续函数（泛函），或 f 在 S 上是上半连续的.

注 2.1.3　由定义 2.1.2 的(1)，f 在点 $\boldsymbol{x}^0$ 处是下半连续的等价于

$$f(\boldsymbol{x}^0) - \varepsilon \leqslant \inf_{\boldsymbol{x} \in U(\boldsymbol{x}^0)} f(\boldsymbol{x}) \leqslant \sup_{U(\boldsymbol{x}^0)} \inf_{\boldsymbol{x} \in U(\boldsymbol{x}^0)} f(\boldsymbol{x}),$$

令 $\varepsilon \to 0$ 即

$$f(\boldsymbol{x}^0) \leqslant \liminf_{\boldsymbol{x} \to \boldsymbol{x}^0} f(\boldsymbol{x}),$$

其中 $\liminf\limits_{\boldsymbol{x} \to \boldsymbol{x}^0} f(\boldsymbol{x}) = \sup\limits_{U(\boldsymbol{x}^0)} \inf\limits_{\boldsymbol{x} \in U(\boldsymbol{x}^0)} f(\boldsymbol{x})$ 是 f 在点 $\boldsymbol{x}^0$ 处的下极限[14]. 同理，由定义 2.1.2 的(2)，f 在点 $\boldsymbol{x}^0$ 处是上半连续的等价于

$$f(\boldsymbol{x}^0) \geqslant \limsup_{\boldsymbol{x} \to \boldsymbol{x}^0} f(\boldsymbol{x}),$$

其中 $\limsup\limits_{\boldsymbol{x} \to \boldsymbol{x}^0} f(\boldsymbol{x}) = \inf\limits_{U(\boldsymbol{x}^0)} \sup\limits_{\boldsymbol{x} \in U(\boldsymbol{x}^0)} f(\boldsymbol{x})$ 是 f 在点 $\boldsymbol{x}^0$ 处的上极限.

定理 2.1.3　设 $\mathscr{X}$ 是线性拓扑空间，$S \subset \mathscr{X}$ 是非空凸集. 若 $f: S \to R$ 是下半连续的严格凸函数，则 f 是 S 上的凸函数.

证明　设对任意的 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in S$，有 $f(\boldsymbol{x}^1) \neq f(\boldsymbol{x}^2)$. 由 f 在 S 上是严格凸的，从(2.1.2)知(2.1.1)成立. 因此，f 在 S 上是凸的.

现设存在 $\bar{\boldsymbol{x}}^1, \bar{\boldsymbol{x}}^2 \in S$，有 $f(\bar{\boldsymbol{x}}^1) = f(\bar{\boldsymbol{x}}^2)$. 用反证法，假设 f 在 S 上不是凸的，则由定义 2.1.1 的(1)知，存在 $\lambda_0 \in (0, 1)$，有

$$f\left(\lambda_0 \bar{\boldsymbol{x}}^1 + (1 - \lambda_0) \bar{\boldsymbol{x}}^2\right) > \lambda_0 f(\bar{\boldsymbol{x}}^1) + (1 - \lambda_0) f(\bar{\boldsymbol{x}}^2)$$

$$= f(\bar{\boldsymbol{x}}^1) = f(\bar{\boldsymbol{x}}^2).$$

记 $\boldsymbol{x}^0 = \lambda_0 \bar{\boldsymbol{x}}^1 + (1-\lambda_0)\bar{\boldsymbol{x}}^2$，由 S 是凸集知 $\boldsymbol{x}^0 \in S$，并从上式得到

$$f(\bar{\boldsymbol{x}}^1) < f(\boldsymbol{x}^0), \quad f(\bar{\boldsymbol{x}}^2) < f(\boldsymbol{x}^0). \tag{2.1.5}$$

对于(2.1.5)的第 1 式，取充分小的 $\varepsilon_0 > 0$，可使 $f(\bar{\boldsymbol{x}}^1) + \varepsilon_0 < f(\boldsymbol{x}^0)$，或

$$f(\bar{\boldsymbol{x}}^1) < f(\boldsymbol{x}^0) - \varepsilon_0.$$

因为 f 在 $\boldsymbol{x}^0 \in S$ 处是下半连续的，由定义 2.1.2 的(1)知，对任意的 $\varepsilon > 0$，存在 λ_0 的邻域 $U(\lambda_0)$ 有

$$-\varepsilon < f[\lambda\bar{\boldsymbol{x}}^1 + (1-\lambda)\bar{\boldsymbol{x}}^2] - f(\boldsymbol{x}^0) \quad \forall \lambda \in U(\lambda_0),$$

从而得

$$f(\bar{\boldsymbol{x}}^1) < f(\boldsymbol{x}^0) - \varepsilon_0 < f[\lambda\bar{\boldsymbol{x}}^1 + (1-\lambda)\bar{\boldsymbol{x}}^2] \quad \forall \lambda \in U(\lambda_0).$$

取 $\lambda' \in U(\lambda_0)$，$\lambda' \neq \lambda_0$，记 $\boldsymbol{x}' = \lambda'\bar{\boldsymbol{x}}^1 + (1-\lambda')\bar{\boldsymbol{x}}^2$，从上式有

$$f(\bar{\boldsymbol{x}}^1) < f(\boldsymbol{x}'). \tag{2.1.6}$$

不妨设 $\boldsymbol{x}^0 \in (\bar{\boldsymbol{x}}^1, \boldsymbol{x}')$，则 $\boldsymbol{x}' \in (\boldsymbol{x}^0, \bar{\boldsymbol{x}}^2)$（若设 $\boldsymbol{x}^0 \in (\boldsymbol{x}', \bar{\boldsymbol{x}}^2)$，则 $\boldsymbol{x}' \in (\bar{\boldsymbol{x}}^1, \boldsymbol{x}^0)$，同理可推证），于是存在 $\lambda_1, \lambda_2 \in (0, 1)$，有 $\boldsymbol{x}^0 = \lambda_1\bar{\boldsymbol{x}}^1 + (1-\lambda_1)\boldsymbol{x}'$ 和 $\boldsymbol{x}' = \lambda_2\boldsymbol{x}^0 + (1-\lambda_2)\bar{\boldsymbol{x}}^2$. 由于 f 在 S 上是严格凸的，依据定义 2.1.1 的(2)并注意到(2.1.6)和(2.1.5)的第 2 式，可得

$$\begin{aligned} f(\boldsymbol{x}^0) &= f[\lambda_1\bar{\boldsymbol{x}}^1 + (1-\lambda_1)\boldsymbol{x}'] \\ &< \lambda_1 f(\bar{\boldsymbol{x}}^1) + (1-\lambda_1)f(\boldsymbol{x}') < f(\boldsymbol{x}') \end{aligned}$$

和

$$\begin{aligned} f(\boldsymbol{x}') &= f[\lambda_2\boldsymbol{x}^0 + (1-\lambda_2)\bar{\boldsymbol{x}}^2] \\ &< \lambda_2 f(\boldsymbol{x}^0) + (1-\lambda_2)f(\bar{\boldsymbol{x}}^2) < f(\boldsymbol{x}^0). \end{aligned}$$

导致以上两式相矛盾. □

为了把凸函数的值域扩展到广义实值域 $R\cup\{\pm\infty\}$，我们对涉及 $\pm\infty$ 的运算作如下约定：

$$\begin{cases}\alpha+\infty=+\infty+\alpha=+\infty \quad \forall\alpha\in R\cup\{+\infty\},\\ \alpha+(-\infty)=-\infty+\alpha=-\infty \quad \forall\alpha\in R\cup\{-\infty\},\\ +\infty+(+\infty)=+\infty,\ -\infty+(-\infty)=-\infty,\\ \lambda\cdot(+\infty)=(+\infty)\cdot\lambda=+\infty,\\ \quad \lambda\cdot(-\infty)=(-\infty)\cdot\lambda=-\infty \quad \forall\lambda>0,\\ \lambda\cdot(+\infty)=(+\infty)\cdot\lambda=-\infty,\\ \quad \lambda\cdot(-\infty)=(-\infty)\cdot\lambda=+\infty \quad \forall\lambda<0,\\ 0\cdot(+\infty)=(+\infty)\cdot 0=0\cdot(-\infty)=(-\infty)\cdot 0=0,\\ +\infty+(+\infty)=+\infty,\\ -(-\infty)=+\infty,\\ \inf\varnothing=+\infty,\ \sup\varnothing=-\infty.\end{cases} \tag{2.1.7}$$

对于(2.1.7)，还规定：只允许使用“消去”而不可使用“移项”. 有了以上规定，我们就可以把任何凸集上的凸函数扩展到全空间.

定义 2.1.3　设 $S\subset\mathscr{Y}$ 是非空凸集，$\hat{f}:S\to R$ 是凸函数. 令

$$f(x)=\begin{cases}\hat{f}(x), & x\in S;\\ +\infty, & x\in\mathscr{Y}\setminus S,\end{cases}$$

则称 $f:\mathscr{Y}\to R\cup\{\pm\infty\}$ 是空间 $\mathscr{Y}$ 上的广义实值凸函数，也简称凸函数.

依照涉及 $+\infty$ 和 $-\infty$ 的约定运算，定义 2.1.3 和定义 2.1.1 的(1)是一致的.

定义 2.1.4　设 $f:\mathscr{Y}\to R\cup\{\pm\infty\}$ 是广义实值函数，则称集合

$$\operatorname{dom}f=\{x\in\mathscr{Y}\mid f(x)<+\infty\}$$

是 f 的有效域. 若 $\operatorname{dom}f\neq\varnothing$，则称 f 是正常函数，或函数 f 是正常的.

定义 2.1.5 设 $f: \mathscr{Y} \to R \cup \{\pm\infty\}$ 是凸函数. 若 $\mathrm{dom}f \neq \varnothing$，并且对于任意的 $\boldsymbol{x} \in \mathrm{dom}f$，有 $f(\boldsymbol{x}) \neq -\infty$，则称 f 是($\mathrm{dom}f$ 上的)正常凸函数，或函数 f 是正常凸的. 否则，称 f 是非正常凸函数. 若 $-f$ 是正常凸函数，则称 f 是正常凹函数.

注 2.1.4 设 $f: \mathscr{Y} \to R \cup \{\pm\infty\}$ 是凸函数，由定义 2.1.3、定义 2.1.1 的(1)和定义 2.1.4，不难得知 $\mathrm{dom}f$ 是凸集.

定理 2.1.4 设 $f_i: \mathscr{Y} \to R \cup \{\pm\infty\}$ $(i=1, 2)$ 是正常凸函数.

(1) $\mathrm{dom}(f_1 + f_2) = \mathrm{dom}f_1 \cap \mathrm{dom}f_2$.

(2) 若 $\mathrm{dom}f_1 \cap \mathrm{dom}f_2 \neq \varnothing$，则 $f_1 + f_2$ 是正常凸函数.

证明 (1) 设 $\boldsymbol{x} \in \mathrm{dom}(f_1 + f_2)$，由定义 2.1.4 有 $f_1(\boldsymbol{x}) + f_2(\boldsymbol{x}) < +\infty$. 又因为 f_1 和 f_2 是正常凸函数，由定义 2.1.5 知 $f_1(\boldsymbol{x})$ 和 $f_2(\boldsymbol{x})$ 均不取 $-\infty$. 由此推知，有 $f_1(\boldsymbol{x}) < +\infty$ 和 $f_2(\boldsymbol{x}) < +\infty$，于是得 $\boldsymbol{x} \in \mathrm{dom}f_1 \cap \mathrm{dom}f_2$. 反之亦然.

(2) 记 $f(\boldsymbol{x}) = f_1(\boldsymbol{x}) + f_2(\boldsymbol{x})$. 任取 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in \mathscr{Y}$ 和任意的 $\lambda \in (0, 1)$，因为 f_1 和 f_2 是凸函数，从而有

$$f_i(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \leqslant \lambda f_i(\boldsymbol{x}^1) + (1-\lambda)f_i(\boldsymbol{x}^2) \quad (i=1, 2).$$

将上式对 i 相加，得

$$f(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \leqslant \lambda f(\boldsymbol{x}^1) + (1-\lambda)f(\boldsymbol{x}^2),$$

因此 f 是凸函数. 从(1)知

$$\mathrm{dom}f = \mathrm{dom}(f_1 + f_2) = \mathrm{dom}f_1 \cap \mathrm{dom}f_2 \neq \varnothing,$$

故对任意的 $\boldsymbol{x} \in \mathrm{dom}f$，有 $\boldsymbol{x} \in \mathrm{dom}f_1$ 和 $\boldsymbol{x} \in \mathrm{dom}f_2$. 由 f_1 和 f_2 是正常凸函数，得 $f_1(\boldsymbol{x}) \neq -\infty$ 和 $f_2(\boldsymbol{x}) \neq -\infty$. 于是，$f = f_1 + f_2$ 是正常凸函数. ▯

现在引进广义实值函数的上图象和下图象，并讨论当它们是凸集时分别与凸函数和凹函数的关系.

定义 2.1.6 设 $f: \mathscr{Y} \to R \cup \{\pm\infty\}$ 是广义实值函数.

(1) 集合

$$\mathrm{epi}f = \{(\boldsymbol{x}, \eta) \in \mathscr{Y} \times R \mid \eta \geqslant f(\boldsymbol{x})\}$$

称为 f 的上图象.

(2) 集合

$$\mathrm{hyp}f = \{(\boldsymbol{x}, \eta) \in \mathscr{Y} \times R \mid \eta \leqslant f(\boldsymbol{x})\}$$

称为 f 的下图象.

图 2.1.2 的(a)和(b)分别是单变量广义实值函数的上图象和下图象的图示.

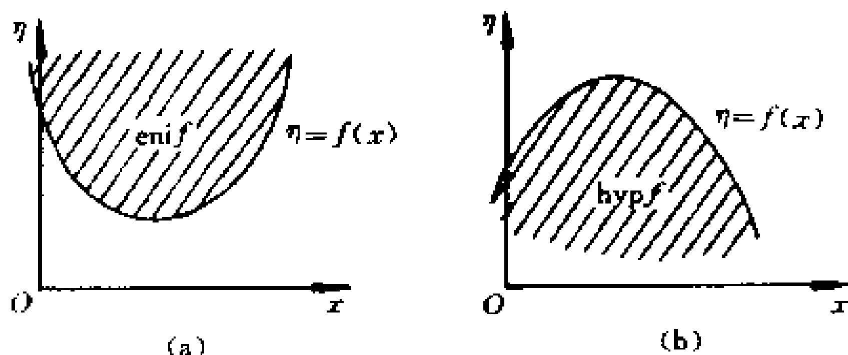


图　2.1.2

定理 2.1.5　设 $f: \mathscr{Y} \to R \cup \{\pm\infty\}$ 是广义实值函数. 以下条件是等价的:

(1) f 是凸函数;

(2) $\mathrm{epi}f$ 是凸集;

(3) $(\mathrm{epi}f)^+ = \{(\boldsymbol{x}, \eta) \in \mathscr{Y} \times R \mid \eta > f(\boldsymbol{x})\}$ 是凸集.

证明　(1)⇒(2). 任取 $(\boldsymbol{x}^1, \eta_1), (\boldsymbol{x}^2, \eta_2) \in \mathrm{epi}f$, 由定义 2.1.6 的(1)有

$$\eta_1 \geqslant f(\boldsymbol{x}^1), \eta_2 \geqslant f(\boldsymbol{x}^2).$$

对于任意的 $\lambda \in (0, 1)$, 因为 f 是凸的, 则得

$$\begin{aligned} f(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) &\leqslant \lambda f(\boldsymbol{x}^1) + (1-\lambda)f(\boldsymbol{x}^2) \\ &\leqslant \lambda\eta_1 + (1-\lambda)\eta_2. \end{aligned}$$

据此，再由定义 2.1.6 的(1) 有

$$\lambda(\boldsymbol{x}^1, \eta_1) + (1-\lambda)(\boldsymbol{x}^2, \eta_2)$$

$$= \left(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2, \lambda\eta_1 + (1-\lambda)\eta_2\right) \in \text{epi}f,$$

于是按定义 1.1.1 知 epi$f$ 是凸集.

(2)⇒(3). 任取 $(\boldsymbol{x}^1, \eta_1), (\boldsymbol{x}^2, \eta_2) \in (\text{epi}f)^+$，则

$$\eta_1 > f(\boldsymbol{x}^1), \ \eta_2 > f(\boldsymbol{x}^2).$$

由此可知，存在 ξ_1、$\xi_2 \in R$，有 $\eta_1 > \xi_1 \geqslant f(\boldsymbol{x}^1)$，$\eta_2 > \xi_2 \geqslant f(\boldsymbol{x}^2)$，从而 $(\boldsymbol{x}^1, \xi_1), (\boldsymbol{x}^2, \xi_2) \in \text{epi}f$. 因为 epi$f$ 是凸集，由定义 1.1.1 知，对任意的 $\lambda \in (0, 1)$ 有

$$(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2, \lambda\xi_1 + (1-\lambda)\xi_2)$$

$$= \lambda(\boldsymbol{x}^1, \xi_1) + (1-\lambda)(\boldsymbol{x}^2, \xi_2) \in \text{epi}f,$$

于是由定义 2.1.6 的(1)得

$$\lambda\eta_1 + (1-\lambda)\eta_2 > \lambda\xi_1 + (1-\lambda)\xi_2 \geqslant f(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2).$$

这说明

$$\lambda(\boldsymbol{x}^1, \eta_1) + (1-\lambda)(\boldsymbol{x}^2, \eta_2)$$

$$= \left(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2, \lambda\eta_1 + (1-\lambda)\eta_2\right) \in (\text{epi}f)^+,$$

于是推得 $(\text{epi}f)^+$ 是凸集.

(3)⇒(1). 任取 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in \mathscr{D}$，则对任意的 $\varepsilon > 0$ 有 $f(\boldsymbol{x}^1) + \varepsilon > f(\boldsymbol{x}^1)$，$f(\boldsymbol{x}^2) + \varepsilon > f(\boldsymbol{x}^2)$，即有 $(\boldsymbol{x}^1, f(\boldsymbol{x}^1) + \varepsilon), (\boldsymbol{x}^2, f(\boldsymbol{x}^2) + \varepsilon) \in (\text{epi}f)^+$. 由于 $(\text{epi}f)^+$ 是凸集，故对任意的 $\lambda \in (0, 1)$ 有

$$\left(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2, \lambda\left(f(\boldsymbol{x}^1) + \varepsilon\right) + (1-\lambda)\left(f(\boldsymbol{x}^2) + \varepsilon\right)\right)$$

$$= \lambda\left(\boldsymbol{x}^1, f(\boldsymbol{x}^1) + \varepsilon\right) + (1-\lambda)\left(\boldsymbol{x}^2, f(\boldsymbol{x}^2) + \varepsilon\right) \in (\text{epi}f)^+,$$

于是有

$$\lambda f(\boldsymbol{x}^1) + (1-\lambda)f(\boldsymbol{x}^2) + \varepsilon > f\big(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2\big).$$

令 $\varepsilon \to 0$，得到

$$f\big(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2\big) \leqslant \lambda f(\boldsymbol{x}^1) + (1-\lambda)f(\boldsymbol{x}^2),$$

因而 f 是凸函数. ▯

注 2.1.5　与定理 2.1.5 类似，f 是凹函数与 $\text{hyp}f$ 是凸集等价.

下面介绍凸函数的有关性质.

定理 2.1.6　设 $f_i: \mathscr{Y} \to R \cup \{\pm\infty\}$ 是凸函数，$\alpha_i \geqslant 0$ $(i = 1, \cdots, m)$. 若

$$f(\boldsymbol{x}) = \sum_{i=1}^{m}\alpha_i f_i(\boldsymbol{x}),$$

则 $f: \mathscr{Y} \to R \cup \{\pm\infty\}$ 是凸函数.

证明　取任意的 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in \mathscr{Y}$ 和任意的 $\lambda \in (0, 1)$，因为 f_i $(i = 1, \cdots, m)$ 是凸函数，由定义 2.1.3 和定义 2.1.1 的(1)，有

$$\begin{aligned}
f\big(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2\big) &= \sum_{i=1}^{m}\alpha_i f_i\big(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2\big) \\
&\leqslant \sum_{i=1}^{m}\alpha_i\big[\lambda f_i(\boldsymbol{x}^1) + (1-\lambda)f_i(\boldsymbol{x}^2)\big] \\
&= \lambda\sum_{i=1}^{m}\alpha_i f_i(\boldsymbol{x}^1) + (1-\lambda)\sum_{i=1}^{m}\alpha_i f_i(\boldsymbol{x}^2) \\
&= \lambda f(\boldsymbol{x}^1) + (1-\lambda)f(\boldsymbol{x}^2).
\end{aligned}$$

据此，即知 f 是凸函数. ▯

定理 2.1.7　设 $f_i: \mathscr{Y} \to R \cup \{\pm\infty\}$ 是凸函数 $(i \in I$，I 是指标集). 若

$$f(\boldsymbol{x}) = \sup_{i\in I} f_i(\boldsymbol{x}),$$

则 $f: \mathscr{Y} \to R \cup \{\pm\infty\}$ 是凸函数.

证明 对任意的 $(x,\eta)\in \mathrm{epi}f$，有 $\eta\geqslant f(x)=\sup\limits_{i\in I}f_i(x)$，它等价于

$$\eta\geqslant f_i(x),\ i\in I.$$

这意味着

$$(x,\eta)\in \mathrm{epi}f_i,\ i\in I,$$

或即 $(x,\eta)\in\bigcap\limits_{i\in I}\mathrm{epi}f_i$，因此 $\mathrm{epi}f=\bigcap\limits_{i\in I}\mathrm{epi}f_i$. 由于 f_i $(i\in I)$ 是凸函数，据定理 2.1.5 知 $\mathrm{epi}f_i$ $(i\in I)$ 是凸集，再由定理 1.1.2 的(2)，得 $\mathrm{epi}f=\bigcap\limits_{i\in I}\mathrm{epi}f_i$ 是凸集. 于是，由定理 2.1.5 即知 f 是凸函数. ▯

注 2.1.6 特别是，在定理 2.1.7 中设各 f_i $(i\in I)$ 是仿射函数，则该定理表示线性空间上仿射函数的上包络是凸函数. 事实上，能够证明线性空间上的正常凸函数在其有效域内部可以用仿射函数的上包络表示，而线性拓扑空间上的下半连续正常凸函数可以用仿射函数的上包络表示[24].

定理 2.1.8 设 $f:\mathscr{X}\to R\cup\{\pm\infty\}$ 和 $g:R\to R\cup\{\pm\infty\}$ 是广义实值函数，并且 $g(t)=f(x+td)$（其中 $d\in\mathscr{X}$），则 f 是凸函数当且仅当对任意的 $d\in\mathscr{X}$，g 是凸函数.

证明 必要性. 对任意的 $d\in\mathscr{X}$，任取 $t_1,t_2\in R$ 和 $\lambda\in(0,1)$，由于 f 是凸函数，有

$$\begin{aligned}g(\lambda t_1+(1-\lambda)t_2)&=f\big(x+[\lambda t_1+(1-\lambda)t_2]d\big)\\&=f\big(\lambda[x+t_1d]+(1-\lambda)[x+t_2d]\big)\\&\leqslant\lambda f(x+t_1d)+(1-\lambda)f(x+t_2d)\\&=\lambda g(t_1)+(1-\lambda)g(t_2).\end{aligned}$$

因此，g 是凸函数.

充分性. 任取 $x^1,x^2\in\mathscr{X}$ 和 $\lambda\in(0,1)$，则有

$$f(\lambda x^1+(1-\lambda)x^2)=f(x^1+[\lambda\cdot 0+(1-\lambda)\cdot 1](x^2-x^1)).$$

令 $x^1=x$，$\lambda\cdot 0+(1-\lambda)\cdot 1=t$，$x^2-x^1=d$. 从上式注意到 g 是凸函数，可得

$$\begin{aligned} f(\lambda x^1+(1-\lambda)x^2) &= g(\lambda\cdot 0+(1-\lambda)\cdot 1) \\ &\leqslant \lambda g(0)+(1-\lambda)g(1) \\ &= \lambda f(x^1)+(1-\lambda)f(x^2), \end{aligned}$$

从而 f 是凸函数. □

定理 2.1.9　设 $h: \mathcal{Y}\to R$ 是凸函数，$g: R\to R\cup\{\pm\infty\}$ 是非减的正常凸函数，$f(x)=g(h(x))$，则 $f: \mathcal{Y}\to R\cup\{\pm\infty\}$ 是凸函数.

证明　任取 $x^1, x^2\in\mathcal{Y}$ 和 $\lambda\in(0,1)$，由于 h 是凸的，g 是非减的，从而有

$$\begin{aligned} f(\lambda x^1+(1-\lambda)x^2) &= g(h(\lambda x^1+(1-\lambda)x^2)) \\ &\leqslant g(\lambda h(x^1)+(1-\lambda)h(x^2)). \end{aligned} \tag{2.1.8}$$

又因 g 是正常凸函数，故有

$$\begin{aligned} g(\lambda h(x^1)+(1-\lambda)h(x^2)) &\leqslant \lambda g(h(x^1))+(1-\lambda)g(h(x^2)) \\ &=\lambda f(x^1)+(1-\lambda)f(x^2). \end{aligned}$$

将上式代入(2.1.8)得到

$$f(\lambda x^1+(1-\lambda)x^2)\leqslant \lambda f(x^1)+(1-\lambda)f(x^2),$$

故 f 是凸函数. □

利用函数上图象的凸包，可以给出函数的凸包概念.

定义 2.1.7　设 $f: \mathcal{Y}\to R\cup\{\pm\infty\}$ 是广义实值函数，令

$$g(x)=\inf\{\eta\mid(x,\eta)\in\mathrm{co}(\mathrm{epi}f)\},$$

则称 $g: \mathscr{X} \to R \cup \{\pm\infty\}$ 是 f 的凸包(函数),记作 $\mathrm{co}(f)$.

为刻画凸包函数的特性,再引进函数的弱函数和弱凸函数概念如下.

定义 2.1.8 设 $f: \mathscr{X} \to R \cup \{\pm\infty\}$ 和 $h: \mathscr{X} \to R \cup \{\pm\infty\}$ 是广义实值函数. 若

$$h(\boldsymbol{x}) \leqslant f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \mathscr{X},$$

则称 h 是 f 的弱函数. 若 h 还是 $\mathscr{X}$ 上的凸函数,则称 h 是 f 的弱凸函数.

下述定理表明,任意函数的凸包是一凸函数.

定理 2.1.10 设 $f: \mathscr{X} \to R \cup \{\pm\infty\}$ 是广义实值函数,则 $\mathrm{co}(f)$ 是 f 的最大弱凸函数.

证明 由 $\mathrm{co}(\mathrm{epi}f)$ 是凸集,从定义 2.1.7 可以推知 $\mathrm{co}(f)$ 是凸函数,并且有

$$\mathrm{epi}f \subset \mathrm{co}(\mathrm{epi}f) \subset \mathrm{epi}(\mathrm{co}(f)).$$

据此,由定义 2.1.6 的(1)得

$$\mathrm{co}(f)(\boldsymbol{x}) \leqslant f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \mathscr{X},$$

因而按定义 2.1.8 知 $\mathrm{co}(f)$ 是 f 的弱凸函数. 现设 $h: \mathscr{X} \to R \cup \{\pm\infty\}$ 是任一上图象包含 $\mathrm{co}(\mathrm{epi}f)$ 的凸函数. 由于

$$h(\boldsymbol{x}) = \inf\{\eta : h(\boldsymbol{x}) \leqslant \eta\} = \inf\{\eta \mid h(\boldsymbol{x}) \in \mathrm{epi}h\},$$

$$\mathrm{co}(f)(\boldsymbol{x}) = \inf\{\eta \mid h(\boldsymbol{x}) \in \mathrm{epi}f\},$$

而 $\mathrm{co}(\mathrm{epi}f) \subset \mathrm{epi}h$, 故有

$$h(\boldsymbol{x}) \leqslant \mathrm{co}(f)(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \mathscr{X}.$$

由 h 的任意性,即得 $\mathrm{co}(f)$ 是 f 的最大弱凸函数. □

最后,对于 Euclid 空间 R^n 上的可微凸函数,我们依次给出它的一阶和二阶的判别条件.

设 $f: R^n \to R$ 一阶可微, $\boldsymbol{x} = (x_1, \cdots, x_n)^T$, 记

$$\nabla f(\boldsymbol{x}) = \left(\frac{\partial f(\boldsymbol{x})}{\partial x_1}, \cdots, \frac{\partial f(\boldsymbol{x})}{\partial x_n} \right)^T$$

是 f 在点 $\boldsymbol{x}$ 处的梯度.

定理 2.1.11 设 $S \subset R^n$ 是开凸集，$f: S \to R$ 一阶可微.

(1) f 是 S 上的凸函数当且仅当

$$\nabla f(\boldsymbol{x}^2)^T(\boldsymbol{x}^1 - \boldsymbol{x}^2) \leqslant f(\boldsymbol{x}^1) - f(\boldsymbol{x}^2) \quad \forall \boldsymbol{x}^1, \boldsymbol{x}^2 \in S. \tag{2.1.9}$$

(2) f 是 S 上的严格(强)凸函数当且仅当

$$\nabla f(\boldsymbol{x}^2)^T(\boldsymbol{x}^1 - \boldsymbol{x}^2) < f(\boldsymbol{x}^1) - f(\boldsymbol{x}^2) \quad \forall \boldsymbol{x}^1, \boldsymbol{x}^2 \in S,$$

$$f(\boldsymbol{x}^1) \neq f(\boldsymbol{x}^2) \ (\boldsymbol{x}^1 \neq \boldsymbol{x}^2). \tag{2.1.10}$$

证明 (1) 必要性. 从 f 是 S 上的凸函数，由定义 2.1.1 的 (1)知，对任意的 $\boldsymbol{x}^1$, $\boldsymbol{x}^2 \in S$ 和任意的 $\lambda \in (0, 1)$，有

$$f(\lambda \boldsymbol{x}^1 + (1 - \lambda)\boldsymbol{x}^2) \leqslant \lambda f(\boldsymbol{x}^1) + (1 - \lambda) f(\boldsymbol{x}^2),$$

从而

$$\frac{f\big(\boldsymbol{x}^2 + \lambda(\boldsymbol{x}^1 - \boldsymbol{x}^2)\big) - f(\boldsymbol{x}^2)}{\lambda} \leqslant f(\boldsymbol{x}^1) - f(\boldsymbol{x}^2).$$

将上式左端的分子作 Taylor 展开，则有

$$\nabla f(\boldsymbol{x}^2)^T(\boldsymbol{x}^1 - \boldsymbol{x}^2) + \frac{o\big(\|\lambda(\boldsymbol{x}^1 - \boldsymbol{x}^2)\|\big)}{\lambda} \leqslant f(\boldsymbol{x}^1) - f(\boldsymbol{x}^2),$$

令 $\lambda \to 0$ 便得(2.1.9).

充分性. 对任意的 $\boldsymbol{x}^1$, $\boldsymbol{x}^2 \in S$ 和任意的 $\lambda \in (0, 1)$，因为 S 是凸集，故由定义 1.1.1 知

$$\boldsymbol{x} = \lambda \boldsymbol{x}^1 + (1 - \lambda)\boldsymbol{x}^2 \in S. \tag{2.1.11}$$

在(2.1.9)中令 $\boldsymbol{x}^1 = \boldsymbol{x}^1$, $\boldsymbol{x}^2 = \boldsymbol{x}$，则有

$$\nabla f(\boldsymbol{x})^T(\boldsymbol{x}^1 - \boldsymbol{x}) \leqslant f(\boldsymbol{x}^1) - f(\boldsymbol{x}) \quad \forall \boldsymbol{x}^1 \in S, \tag{2.1.12}$$

再在(2.1.9)中令 $\boldsymbol{x}^1=\boldsymbol{x}^2$, $\boldsymbol{x}^2=\boldsymbol{x}$, 又有

$$\nabla f(\boldsymbol{x})^T(\boldsymbol{x}^2-\boldsymbol{x})\leqslant f(\boldsymbol{x}^2)-f(\boldsymbol{x})\quad \forall \boldsymbol{x}^2\in S.\tag{2.1.13}$$

作 λ(2.1.12) + $(1-\lambda)$(2.1.13), 得到

$$\nabla f(\boldsymbol{x})^T\left[\lambda\boldsymbol{x}^1+(1-\lambda)\boldsymbol{x}^2-\boldsymbol{x}\right]$$
$$\leqslant \lambda f(\boldsymbol{x}^1)+(1-\lambda)f(\boldsymbol{x}^2)-f(\boldsymbol{x}).$$

注意到(2.1.11),便有

$$f\left(\lambda\boldsymbol{x}^1+(1-\lambda)\boldsymbol{x}^2\right)=f(\boldsymbol{x})\leqslant\lambda f(\boldsymbol{x}^1)+(1-\lambda)f(\boldsymbol{x}^2),$$

于是由定义 2.1.1 的(1)得 f 是 S 上的凸函数.

(2) 将(1)证明中的不等式改为严格不等式,类似地可以推证. □

设 $f:R^n\to R$ 二阶可微, $\boldsymbol{x}=(x_1,\cdots,x_n)^T$, 记

$$\nabla^2 f(\boldsymbol{x})=\begin{pmatrix}\dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_1^2} & \cdots & \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_1\partial x_n}\\ \cdots & \cdots & \cdots\\ \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_n\partial x_1} & \cdots & \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_n^2}\end{pmatrix}$$

是 f 在点 $\boldsymbol{x}$ 处的 Hesse 矩阵.

定理 2.1.12 设 $S\subset R^n$ 是开凸集, $f: S\to R$ 二阶连续可微.

(1) f 是 S 上的凸函数当且仅当对任意的 $\boldsymbol{x}\in S$, $\nabla^2 f(\boldsymbol{x})$ 是半正定的.

(2) 若对任意的 $\boldsymbol{x}\in S$, $\nabla^2 f(\boldsymbol{x})$ 是正定的,则 f 是 S 上的强凸函数.

证明 (1) 充分性. 对任意的 $\boldsymbol{x}^1,\boldsymbol{x}^2\in S$, 有 Taylor 公式

$$f(\boldsymbol{x}^1)=f(\boldsymbol{x}^2)+\nabla f(\boldsymbol{x}^2)^T(\boldsymbol{x}^1-\boldsymbol{x}^2)$$
$$+\frac{1}{2}(\boldsymbol{x}^1-\boldsymbol{x}^2)^T\nabla^2 f(\boldsymbol{\xi})(\boldsymbol{x}^1-\boldsymbol{x}^2),$$

其中$\xi \in S$. 由于$\nabla^2 f(\xi)$是半正定的,故$(x^1 - x^2)^T \nabla^2 f(\xi)(x^1 - x^2) \geqslant 0$, 于是从上式推知

$$\nabla f(x^2)^T(x^1 - x^2) \leqslant f(x^1) - f(x^2).$$

据此,由定理2.1.11的(1)得知f在S上是凸的.

必要性. 任取$x \in S$, 因为S是开集,故对任意的$y \in R^n$, 存在充分小的$\lambda > 0$有$x + \lambda y \in S$. 由于f是S上的凸函数,在定理2.1.11中的(1)中令$x^1 = x + \lambda y$, $x^2 = x$, 则得

$$f(x) + \lambda \nabla f(x)^T y \leqslant f(x + \lambda y). \qquad (2.1.14)$$

另外,由Taylor公式有

$$f(x + \lambda y) = f(x) + \lambda \nabla f(x)^T y + \frac{1}{2}\lambda^2 y^T \nabla^2 f(\xi) y,$$

其中$\xi = x + \theta\lambda y$ ($\theta \in (0, 1)$). 将它与(2.1.14)作比较,得知$y^T \nabla^2 f(\xi) y \geqslant 0$. 令$\lambda \to 0$,则$\xi \to x$, 于是得到$y^T \nabla^2 f(x) y \geqslant 0$, 即$\nabla^2 f(x)$是半正定的.

(2) 对任意的$x^1, x^2 \in S$, $x^1 \neq x^2$, 由Taylor公式有

$$f(x^1) = f(x^2) + \nabla f(x^2)^T(x^1 - x^2) + \frac{1}{2}(x^1 - x^2)^T \nabla^2 f(\xi)(x^1 - x^2), \qquad (2.1.15)$$

其中$\xi \in S$. 因对任意的$x \in S$, $\nabla^2 f(x)$是正定的,故$(x^1 - x^2)^T \nabla^2 f(\xi)(x^1 - x^2) > 0$. 注意到$x^1 \neq x^2$, 从(2.1.15)推知

$$\nabla f(x^2)^T(x^1 - x^2) < f(x^1) - f(x^2).$$

据此,由定理2.1.11的(2)得到f在S上是强凸的. □

§2.2 凸函数的连续性和对偶性

连续性是函数的一个重要特性. 在这一节,先讨论线性拓扑空

间上凸函数的连续性问题.然后研究线性拓扑空间上的凸函数与对偶空间上相应的共轭函数之间的对偶关系.

设 $\mathscr{X}$ 是线性拓扑空间.

定义 2.2.1　设 $f:\mathscr{X}\to R\cup\{\pm\infty\}$ 是广义实值函数,点 $x^0\in\mathscr{X}$. 若对任意的 $\varepsilon>0$,存在点 x^0 的邻域 $U(x^0)$,有

$$|f(x)-f(x^0)|<\varepsilon\quad\forall x\in U(x^0),$$

则称 f 在点 x^0 处是连续的.若 f 在集合 $S\subset\mathscr{X}$ 的每一点处都是连续的,则称 f 是 S 上的**连续函数(连续泛函)**,或 f 在 S 上是连续的.

注 2.2.1　由定义 2.2.1 和定义 2.1.2 易知,若 f 在点 $x^0\in\mathscr{X}$ 处是连续的,则 f 在点 x^0 处既是下半连续的又是上半连续的;反之,若 f 在点 x^0 处既是下半连续的又是上半连续的,则 f 在点 x^0 处是连续的.

定理 2.2.1　设 $f:\mathscr{X}\to R\cup\{+\infty\}$ 是正常凸函数.若存在点 $x^0\in\operatorname{int}(\operatorname{dom}f)$ 使得 f 在 x^0 的某邻域内上有界,则 f 在 $\operatorname{int}(\operatorname{dom}f)$ 内是连续的.

证明　不失一般性,我们设 $x^0=\boldsymbol{0}$(否则,可令 $y=x-x^0$ 将 $f(x)$ 转为对 $g(y)=f(y+x^0)$ 的讨论). 由已设 f 在 $\boldsymbol{0}$ 的邻域内上有界,则存在点 $\boldsymbol{0}$ 的对称邻域 $U(\boldsymbol{0})$ 和常数 $M\in R$,有

$$f(x)\leqslant M\quad\forall x\in U(\boldsymbol{0}).\tag{2.2.1}$$

先证 f 在点 $\boldsymbol{0}$ 处是连续的.为此,任给 $\varepsilon\in(0,1)$,则当 $x\in\varepsilon U(\boldsymbol{0})$ 时有 $\pm\dfrac{x}{\varepsilon}\in U(\boldsymbol{0})$,从而根据 f 是 $U(\boldsymbol{0})$ 上的凸函数,按定义 2.1.1 的(1)有

$$f(x)=f\left(\varepsilon\frac{x}{\varepsilon}+(1-\varepsilon)\cdot\boldsymbol{0}\right)\leqslant\varepsilon f\left(\frac{x}{\varepsilon}\right)+(1-\varepsilon)f(\boldsymbol{0}).$$

注意到(2.2.1),有 $f(x)\leqslant\varepsilon M+(1-\varepsilon)f(\boldsymbol{0})$,故得

$$f(x)-f(\boldsymbol{0})\leqslant\varepsilon[M-f(\boldsymbol{0})].\tag{2.2.2}$$

另外，由 f 是凸函数和(2.2.1)又有

$$f(\mathbf{0})=f\left(\frac{1}{1+\varepsilon}\mathbf{x}+\frac{\varepsilon}{1+\varepsilon}\left(\frac{-\mathbf{x}}{\varepsilon}\right)\right)$$

$$\leqslant\frac{1}{1+\varepsilon}f(\mathbf{x})+\frac{\varepsilon}{1+\varepsilon}f\left(\frac{-\mathbf{x}}{\varepsilon}\right)$$

$$\leqslant\frac{1}{1+\varepsilon}f(\mathbf{x})+\frac{\varepsilon}{1+\varepsilon}M,$$

于是

$$f(\mathbf{x})-f(\mathbf{0})\geqslant-\varepsilon[M-f(\mathbf{0})].\qquad(2.2.3)$$

由(2.2.2)和(2.2.3)得到

$$\left|f(\mathbf{x})-f(\mathbf{0})\right|\leqslant\varepsilon[M-f(\mathbf{0})],$$

因为 $\varepsilon>0$ 是任给的，所以由定义 2.2.1 知 f 在点 $\mathbf{0}$ 处是连续的.

现任取 $\mathbf{x}\in\operatorname{int}(\operatorname{dom}f)$，我们证明 f 在点 $\mathbf{x}$ 处是连续的. 事实上，这时存在 $\rho>1$ 使得 $\rho\mathbf{x}\in\operatorname{dom}f$. 因为 $U(\mathbf{0})$ 是点 $\mathbf{0}$ 的邻域，故

$$U(\mathbf{x})=\mathbf{x}+\left(1-\frac{1}{\rho}\right)U(\mathbf{0})$$

是点 $\mathbf{x}$ 的邻域. 由此，对任意的 $\mathbf{y}\in U(\mathbf{x})$，存在 $\mathbf{u}\in U(\mathbf{0})$ 有

$$\mathbf{y}=\mathbf{x}+\left(1-\frac{1}{\rho}\right)\mathbf{u}.$$

根据 f 在 $\operatorname{dom}f$ 上是凸的以及(2.2.1)，我们得到

$$f(\mathbf{y})=f\left(\frac{1}{\rho}(\rho\mathbf{x})+\left(1-\frac{1}{\rho}\right)\mathbf{u}\right)$$

$$\leqslant\frac{1}{\rho}f(\rho\mathbf{x})+\left(1-\frac{1}{\rho}\right)f(\mathbf{u})$$

$$\leqslant\frac{1}{\rho}f(\rho\mathbf{x})+\left(1-\frac{1}{\rho}\right)M.$$

由 $\mathbf{y}\in U(\mathbf{x})$ 的任意性，即知 f 在点 $\mathbf{x}$ 的邻域上有界. 于是，与前面

已证得的结果同理,可得 f 在点 $\boldsymbol{x}$ 处是连续的. ▯

推论 2.2.2 设 $f: \mathscr{X} \to R \cup \{+\infty\}$ 是正常凸函数,则 f 在 $\mathrm{int}(\mathrm{dom} f)$ 内是连续的当且仅当 $\mathrm{int}(\mathrm{epi} f) \neq \varnothing$.

证明 因为 $\mathrm{int}(\mathrm{epi} f) \neq \varnothing$ 等价于 f 在某些点的邻域内上有界,所以由定理 2.2.1 并注意到连续函数总是上有界的,它也等价于 f 在 $\mathrm{int}(\mathrm{dom} f)$ 内是连续的. ▯

对于有限维 Euclid 空间 R^n 上的正常凸函数,则由定理 2.2.1 可以推知它在有效域内部一定是连续的.

推论 2.2.3 设 $f: R^n \to R \cup \{+\infty\}$ 是正常凸函数,则 f 在 $\mathrm{int}(\mathrm{dom} f)$ 内是连续的.

证明 在 $\mathrm{int}(\mathrm{dom} f)$ 内可以找到 $n+1$ 个仿射无关的点 $\boldsymbol{x}^0$, $\boldsymbol{x}^1, \cdots, \boldsymbol{x}^n$,以它们为顶点构成的单纯形 $A \subset \mathrm{int}(\mathrm{dom} f)$,并且由 f 的凸性有

$$f(\boldsymbol{x}) = f\Big(\sum_{i=0}^{n} \lambda_i \boldsymbol{x}^i\Big) \leqslant \sum_{i=0}^{n} \lambda_i f(\boldsymbol{x}^i) \quad \forall \boldsymbol{x} = \sum_{i=0}^{n} \lambda_i \boldsymbol{x}^i \in A,$$

其中 $\lambda_i \geqslant 0\ (i = 0, 1, \cdots, n)$, $\sum\limits_{i=0}^{n} \lambda_i = 1$. 据此,$f$ 在 A 的内部上有界,而 A 的内部是 R^n 中的开集,故由定理 2.2.1 得证. ▯

但是,对于无限维空间来说,上述结论一般不成立. 事实上,在线性拓扑空间上可以存在处处有限而处处不连续的凸函数.

下面考虑半连续情况.

定理 2.2.4 设 $f: \mathscr{X} \to R \cup \{+\infty\}$ 是正常凸函数. 若存在点 $\boldsymbol{x}^0 \in \mathrm{dom} f$ 使得 f 在 $\boldsymbol{x}^0$ 处是上半连续的,则 f 在 $\mathrm{int}(\mathrm{dom} f)$ 内是连续的.

证明 因为 f 在点 $\boldsymbol{x}^0$ 处是上半连续的,由定义 2.1.2 的(2)可知 f 在 $\boldsymbol{x}^0$ 的某邻域内上有界. 于是,根据定理 2.2.1 即得结论. ▯

函数的半连续性与相应集合的闭性之间有如下关系.

定理 2.2.5 设 $f: \mathscr{X} \to R \cup \{\pm\infty\}$ 是广义实值函数,则以下条件是等价的:

(1) f 是下半连续函数；

(2) $\mathrm{epi}f \subset \mathscr{H} \times R$ 是闭集；

(3) 对任一 $c \in R$，$\{x \in \mathscr{H} \mid f(x) \leqslant c\}$ 是闭集；

(4) 对任一 $c \in R$，$\{x \in \mathscr{H} \mid f(x) > c\}$ 是开集.

证明　(1)⇒(2). 任取点列 $\{(x^k, \eta_k)\} \subset \mathrm{epi}f$，并且 $(x^k, \eta_k) \to (x^0, \eta_0)$ $(k \to \infty)$. 因为 $(x^k, \eta_k) \in \mathrm{epi}f (k = 1, 2, \cdots)$，所以由定义 2.1.6 的(1)有

$$f(x^k) \leqslant \eta_k, k = 1, 2, \cdots.$$

由于 f 在点 x^0 处是下半连续的，于是由定义 2.1.2 的(1)可知，对任给的 $\varepsilon > 0$，存在 x^0 的邻域 $U(x^0)$ 有

$$f(x^0) - \varepsilon \leqslant \inf_{x \in U(x^0)} f(x).$$

令 $\varepsilon \to 0$，特别地有

$$f(x^0) \leqslant \lim_{k \to \infty} \inf f(x^k) \leqslant \lim_{k \to \infty} \inf \eta_k = \eta_0.$$

据此，再按定义 2.1.6 的(1)即得 $(x^0, \eta_0) \in \mathrm{epi}f$，故知 $\mathrm{epi}f$ 是闭集.

(2)⇒(3). 任取点列 $\{x^k\} \subset \{x \in \mathscr{H} \mid f(x) \leqslant c\}$，并且 $x^k \to x^0$ $(k \to \infty)$. 因为 $f(x^k) \leqslant c$ $(k = 1, 2, \cdots)$，由定义 2.1.6 的(1)有 $(x^k, c) \in \mathrm{epi}f$. 由于 $\mathrm{epi}f$ 是闭集，因此 $f(x^0) \leqslant c$. 于是有 $x^0 \in \{x \in \mathscr{H} \mid f(x) \leqslant c\}$，从而得证.

(3)⇒(4). 由 $\{x \in \mathscr{H} \mid f(x) \leqslant c\}$ 是闭集，即知 $\{x \in \mathscr{H} \mid f(x) > c\} = \mathscr{H} \setminus \{x \in \mathscr{H} \mid f(x) \leqslant c\}$ 是开集.

(4)⇒(1). 任取 $x^0 \in \mathscr{H}$，对任给的 $\varepsilon > 0$ 令

$$c = f(x^0) - \varepsilon, \tag{2.2.4}$$

则 $f(x^0) = c + \varepsilon > c$，于是 $x^0 \in \{x \in \mathscr{H} \mid f(x) > c\}$. 因为 $\{x \in \mathscr{H} \mid f(x) > c\}$ 是开集，故存在点 x^0 的邻域 $U(x^0)$，有 $U(x^0) \subset \{x \in \mathscr{H} \mid f(x) > c\}$. 据此，有

$$f(\boldsymbol{x}) > c \quad \forall \boldsymbol{x} \in U(\boldsymbol{x}^0),$$

或由(2.2.4)得

$$-\varepsilon < f(\boldsymbol{x}) - f(\boldsymbol{x}^0) \quad \forall \boldsymbol{x} \in U(\boldsymbol{x}^0).$$

按定义 2.1.2 的(1),即知 f 在 $\boldsymbol{x}^0$ 处是下半连续的. 因为 $\boldsymbol{x}^0 \in \mathscr{H}$ 是任取的,故 f 是下半连续函数. ▯

定理 2.2.6 设 $f: \mathscr{H} \to R \cup \{\pm\infty\}$ 是广义实值函数,则以下条件是等价的:

(1) f 是上半连续函数;

(2) $\mathrm{hpy} f \subset \mathscr{H} \times R$ 是闭集;

(3) 对任一 $c \in R$, $\{\boldsymbol{x} \in \mathscr{H} \mid f(\boldsymbol{x}) \geqslant c\}$ 是闭集;

(4) 对任一 $c \in R$, $\{\boldsymbol{x} \in \mathscr{H} \mid f(\boldsymbol{x}) < c\}$ 是开集.

证明 与定理 2.2.5 的证明类似,在此从略. ▯

定理 2.2.7 设 $\mathscr{B}$ 是 Banach 空间, $f: \mathscr{B} \to R \cup \{+\infty\}$ 是正常凸函数. 若 f 在 $\mathscr{B}$ 上是下半连续的,则 f 在 $\mathrm{int}(\mathrm{dom} f)$ 内是连续的.

证明 不妨取 $\boldsymbol{0} \in \mathrm{int}(\mathrm{dom} f)$, 并且 $f(\boldsymbol{0}) < c$. 记

$$A = \{\boldsymbol{x} \in \mathscr{B} \mid f(\boldsymbol{x}) \leqslant c\},$$

因为 f 在 $\mathscr{B}$ 上是下半连续的,由定理 2.2.5 的(1)和(3)知 A 是闭的. 又由 f 的凸性,根据定理 2.1.3 知 A 是凸集. 对任意的 $\boldsymbol{x} \in \mathscr{B}$, 因为 f 是 $\mathscr{B}$ 上的凸函数,故存在 $\bar{\varepsilon} > 0$, 对任意充分小的 $\varepsilon \in (0, \bar{\varepsilon})$ 有

$$f\left(\frac{\varepsilon \boldsymbol{x}}{1+\varepsilon}\right) \leqslant \frac{1}{1+\varepsilon} f(\boldsymbol{0}) + \frac{\varepsilon}{1+\varepsilon} f(\boldsymbol{x}) \leqslant c.$$

由此,有

$$\frac{\varepsilon}{1+\varepsilon} \boldsymbol{x} \in A, \quad 或 \quad \boldsymbol{x} \in \frac{1+\varepsilon}{\varepsilon} A.$$

这说明 A 是吸收集,从而 $U = A \cap (-A)$ 是对称吸收闭凸集. 现在证明 $\mathrm{int} U \neq \varnothing$. 事实上,假若 $\mathrm{int} U = \varnothing$, 则有

$$\text{int}(kU) = \varnothing,\ k = 1, 2, \cdots. \tag{2.2.5}$$

由(2.2.5)，当 $k=1$ 时知，任取 $\boldsymbol{x}^1 \notin U$，则存在 $\delta_1 > 0$，使对 $\boldsymbol{x}^1$ 的 δ_1-邻域 $U(\boldsymbol{x}^1, \delta_1)$ 有

$$U \cap U(\boldsymbol{x}^1, \delta_1) = \varnothing.$$

由(2.2.5)，当 $k = 2$ 时，又知 $U(\boldsymbol{x}^1, \delta_1) \not\subset 2U$. 于是，取 $\boldsymbol{x}^2 \in U(\boldsymbol{x}^1, \delta_1)$ 和 $\boldsymbol{x}^2 \notin 2U$，则存在 $\delta_2 \in \left(0, \frac{\delta_1}{2}\right]$，使对 $\boldsymbol{x}^2$ 的 δ_2-邻域 $U(\boldsymbol{x}^2, \delta_2)$ 有

$$2U \cap U(\boldsymbol{x}^2, \delta_2) = \varnothing.$$

同样地，由(2.2.5)，当 $k=3$ 时，则知 $U(\boldsymbol{x}^2, \delta_2) \not\subset 3U$，故取 $\boldsymbol{x}^3 \in U(\boldsymbol{x}^2, \delta_2)$ 和 $\boldsymbol{x}^3 \notin 3U$，存在 $\delta_3 \in \left(0, \frac{\delta_2}{2}\right]$，使对 $\boldsymbol{x}^3$ 的 δ_3-邻域 $U(\boldsymbol{x}^3, \delta_3)$ 有

$$3U \cap U(\boldsymbol{x}^3, \delta_3) = \varnothing.$$

如此继续下去，可得到序列 $\{\boldsymbol{x}^k\}$ 和正数列 $\{\delta_k\}$，满足 $\boldsymbol{x}^{k+1} \in U(\boldsymbol{x}^k, \delta_k)$ 和 $\delta_{k+1} \in \left(0, \frac{\delta_k}{2}\right]$ $(k = 1, 2, \cdots)$，并且有

$$kU \cap U(\boldsymbol{x}^k, \delta_k) = \varnothing. \tag{2.2.6}$$

显然，$\{\boldsymbol{x}^k\}$ 是 Cauchy 序列，由于 $\mathscr{B}$ 是完备空间，则有 $\boldsymbol{x}^k \to \boldsymbol{x}^* \in \mathscr{B}$. 据此，由(2.2.6)得知 $\boldsymbol{x}^* \notin kU$ $(k = 1, 2, \cdots)$，这与 U 是吸收集矛盾.

从 $\text{int}U \neq \varnothing$，设 $\boldsymbol{x}^0 \in \text{int}U$，则存在 $\delta > 0$ 对于 $\boldsymbol{x}^0$ 的 δ-邻域 $U(\boldsymbol{x}^0, \delta)$ 有 $U(\boldsymbol{x}^0, \delta) \subset U$. 因为 U 是对称凸集，故有 $-\boldsymbol{x}^0 \in U$. 设 $U(\boldsymbol{0}, \delta)$ 是点 $\boldsymbol{0}$ 的 δ-邻域，则有

$$\frac{1}{2}U(\boldsymbol{0}, \delta) = \frac{1}{2}\left[U(\boldsymbol{x}^0, \delta) - \boldsymbol{x}^0\right] \subset \frac{1}{2}U + \frac{1}{2}U = U.$$

显然，f 在 $\frac{1}{2}U(\boldsymbol{0}, \delta)$ 内是上有界的，于是 f 在点 $\boldsymbol{0}$ 的邻域 U 内上

有界. 据此,由定理 2.2.1 得知 f 在 $\mathrm{int}(\mathrm{dom}f)$ 内是连续的. □

定义 2.2.2 设 $f:\mathscr{X}\to R\cup\{\pm\infty\}$, $g:\mathscr{X}\to R\cup\{\pm\infty\}$ 是广义实值函数. 若

$$\mathrm{epi}g=\mathrm{cl}(\mathrm{epi}f),$$

则称 g 是 f 的下半连续包.

定义 2.2.3 设 $f:\mathscr{X}\to R\cup\{+\infty\}$ 是正常凸函数.

(1) 若 $g:\mathscr{X}\to R\cup\{\pm\infty\}$ 是 f 的下半连续包,则称 g 是 f 的闭包(函数),记作 $g=\mathrm{cl}(f)$.

(2) 若 $f=\mathrm{cl}(f)$,则称 f 是闭凸函数,或 f 是闭凸的.

注 2.2.2 设 $f:\mathscr{X}\to R\cup\{+\infty\}$ 是正常凸函数,则由定义 2.2.2 和定义 2.2.3 的(1)可知

$$\mathrm{epi}(\mathrm{cl}(f))=\mathrm{cl}(\mathrm{epi}f). \tag{2.2.7}$$

定理 2.2.8 设 $f:\mathscr{X}\to R\cup\{+\infty\}$ 是正常凸函数,则 f 是 $\mathscr{X}$ 上的下半连续函数当且仅当 f 是闭凸函数.

证明 充分性. 从 f 是闭凸函数,由定义 2.2.3 的(2)有 $f=\mathrm{cl}(f)$. 据此,由(2.2.7)可知 $\mathrm{epi}f=\mathrm{epi}(\mathrm{cl}(f))=\mathrm{cl}(\mathrm{epi}f)$, 因而 $\mathrm{epi}f$ 是闭集. 于是,由定理 2.2.5 的(2)和(1),得知 f 是下半连续的.

必要性. 设 f 是下半连续的,则由定理 2.2.5 知 $\mathrm{epi}f$ 是闭集. 因此,由(2.2.7)有 $\mathrm{epi}f=\mathrm{cl}(\mathrm{epi}f)=\mathrm{epi}(\mathrm{cl}(f))$. 这说明 $f=\mathrm{cl}(f)$, 从而由定义 2.2.3 的(2)得 f 是闭凸函数. □

以下建立线性拓扑空间上的凸函数与对偶空间上相应的共轭函数间的对偶关系.

定义 2.2.4 设 $f:\mathscr{X}\to R\cup\{\pm\infty\}$ 是凸函数, $\mathscr{X}^*$ 是 $\mathscr{X}$ 的对偶空间.

(1) 令

$$f^*(x^*)=\sup_{x\in\mathscr{X}}\{\langle x^*,x\rangle-f(x)\},\quad x^*\in\mathscr{X}^*, \tag{2.2.8}$$

则称 $f^*:\mathscr{X}^*\to R\cup\{\pm\infty\}$ 是 f 的共轭函数.

(2) 令

$$f^{**}(x)=\sup_{x^*\in\mathscr{X}^*}\{\langle x^*, x\rangle - f^*(x^*)\}, x\in\mathscr{X}, \tag{2.2.9}$$

则称 $f^{**}: \mathscr{X}\to R\cup\{\pm\infty\}$ 是 f 的二次共轭函数.

通常,称由(2.2.8)所确定的 f 与 f^* 之间的关系具有共轭关系或对偶关系. 显然,由定义 2.2.4,即知有 $(f^*)^*=f^{**}$.

注 2.2.3　从定义 2.2.4 的(1),可以直接得到如下的 Young-Fenchel 不等式:

$$\langle x^*, x\rangle \leqslant f^*(x^*)+f(x)\quad \forall x^*\in\mathscr{X}^*, \forall x\in\mathscr{X}. \tag{2.2.10}$$

注 2.2.4　由定义 2.2.4 的(1)可知,若存在某个 $\bar{x}\in\mathscr{X}$ 使得 $f(\bar{x})=-\infty$,则有 $f^*(x^*)\equiv+\infty$;反之,若存在某个 $\bar{x}^*\in\mathscr{X}^*$ 使得 $f^*(\bar{x}^*)=-\infty$,则仅当 $f(x)\equiv+\infty$ 时才有可能,从而使 $f^*(x^*)\equiv-\infty$. 因此,当 $f\equiv+\infty$ 或 f 取 $-\infty$ 的情形都是没有意义的. 此后,我们只考虑 $f: \mathscr{X}\to R\cup\{+\infty\}$ 并且 $f\not\equiv+\infty$ 的情形.

以下介绍共轭函数的有关性质.

定理 2.2.9　设 $f: \mathscr{X}\to R\cup\{+\infty\}$ 是凸函数,则

$$\inf_{x\in\mathscr{X}} f(x)=-f^*(\mathbf{0}).$$

证明　由定义 2.2.4 的(1)直接可知 $f^*(\mathbf{0})=\sup_{x\in\mathscr{X}}\{\langle \mathbf{0}, x\rangle - f(x)\}=-\inf_{x\in\mathscr{X}} f(x)$. □

定理 2.2.10　设 $f: \mathscr{X}\to R\cup\{+\infty\}$ 是凸函数.

(1) $f^*: \mathscr{X}^*\to R\cup\{+\infty\}$ 是下半连续正常凸函数.

(2) $f^{**}: \mathscr{X}\to R\cup\{+\infty\}$ 是下半连续正常凸函数.

证明　(1) 为证 f^* 是下半连续的,由定理 2.2.5,我们证明 $\operatorname{epi} f^*$ 是闭集. 事实上,取点列 $\{(x^k, \eta_k)\}\subset \operatorname{epi} f^*$,并且 $(x^k, \eta_k)\to(x^0, \eta_0)$ $(k\to\infty)$,因为 $(x^k, \eta_k)\in\operatorname{epi} f^*$ $(k=1, 2, \cdots)$,所以由定义 2.1.6 的(1)和定义 2.2.4 的(1)有

$$\eta_k \geqslant f^*(x^k) = \sup_{x\in\mathscr{X}}\{\langle x^k, x\rangle - f(x)\},\ x^k \in \mathscr{X}^*,\ k=1,2,\cdots.$$

令 $k\to\infty$，则得

$$\eta_0 = \lim_{k\to\infty}\eta_k \geqslant \lim_{k\to\infty}\sup_{x\in\mathscr{X}}\{\langle x^k, x\rangle - f(x)\}$$

$$= \sup_{x\in\mathscr{X}}\{\langle x^0, x\rangle - f(x)\} = f^*(x^0).$$

由此，我们有 $(x^0, \eta_0)\in \mathrm{epi}f^*$，故 $\mathrm{epi}f^*$ 是闭的.

现证 f^* 是凸的. 为此，任取 $x^1, x^2\in\mathscr{X}^*$ 和 $\lambda\in(0,1)$，由定义 2.2.4 的(1)有

$$f^*(\lambda x^1 + (1-\lambda)x^2) = \sup_{x\in\mathscr{X}}\left\{\langle \lambda x^1 + (1-\lambda)x^2, x\rangle - f(x)\right\}$$

$$\leqslant \lambda\sup_{x\in\mathscr{X}}\left\{\langle x^1, x\rangle - f(x)\right\}$$

$$+ (1-\lambda)\sup_{x\in\mathscr{X}}\left\{\langle x^2, x\rangle - f(x)\right\}$$

$$= \lambda f^*(x^1) + (1-\lambda)f^*(x^2).$$

于是，根据定义 2.1.1 的(1)和定义 2.1.3，即知 f^* 是凸函数. 由注 2.2.4，因为 $f^*\neq -\infty$，故 f^* 是正常凸函数.

(2) 与(1)的证明类似. ▯

定理 2.2.11 设 $f:\mathscr{X}\to R\cup\{+\infty\}$，$h:\mathscr{X}\to R\cup\{+\infty\}$ 是凸函数. 若 h 是 f 的弱函数，则 f^* 是 h^* 的弱函数.

证明 从 h 是 f 的弱函数，由定义 2.1.8 有

$$h(x)\leqslant f(x)\quad \forall x\in\mathscr{X}.$$

据此，按定义 2.2.4 的(1)，即知对任意的 $x^*\in\mathscr{X}^*$，有

$$f^*(x^*) = \sup_{x\in\mathscr{X}}\left\{\langle x^*, x\rangle - f(x)\right\}$$

$$\leqslant \sup_{x\in\mathscr{X}}\left\{\langle x^*, x\rangle - h(x)\right\}$$

$$= h^*(x^*),$$

故 f^* 是 h^* 的弱函数. ▯

定理 2.2.12 设 $f: \mathscr{X} \to R \cup \{+\infty\}$ 是凸函数.

(1) f^{**} 是 f 的弱函数，即对任意的 $x \in \mathscr{X}$ 有 $f^{**}(\boldsymbol{x}) \leqslant f(\boldsymbol{x})$.

(2) $f^{**} = f$ 当且仅当 f 是下半连续正常凸函数.

(3) $f^{***} = f^{*}$.

证明 (1) 由(2.2.10)(Young-Fenchel 不等式)有

$$\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle - f^*(\boldsymbol{x}^*) \leqslant f(\boldsymbol{x}) \quad \forall \boldsymbol{x}^* \in \mathscr{X}^*, \forall \boldsymbol{x} \in \mathscr{X}.$$

于是，由定义 2.2.4 的(2)，即得

$$f^{**}(\boldsymbol{x}) = \sup_{\boldsymbol{x}^* \in \mathscr{X}^*} \left\{ \langle \boldsymbol{x}^*, \boldsymbol{x} \rangle - f^*(\boldsymbol{x}^*) \right\} \leqslant f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \mathscr{X}.$$

(2) 由定理 2.2.10 中的(2)知，$f = f^{**}$ 是下半连续正常凸函数. 反之，设 f 是下半连续正常凸函数，则按注 2.1.6，它可以用仿射函数的上包络表示. 于是，存在 $\{\boldsymbol{x}_i^*\}_{i \in I} \subset \mathscr{X}^*$ 和 $\{\alpha_i\}_{i \in I} \in R$ (I 是指标集)，使得

$$f(\boldsymbol{x}) = \sup_{i \in I} \{ \langle \boldsymbol{x}_i^*, \boldsymbol{x} \rangle - \alpha_i \}. \tag{2.2.11}$$

据此，由定义 2.2.4 的(1)知，对 $\boldsymbol{x}^* \in \mathscr{X}^*$ 有

$$\begin{aligned} f^*(\boldsymbol{x}^*) &= \sup_{\boldsymbol{x} \in \mathscr{X}} \left\{ \langle \boldsymbol{x}^*, \boldsymbol{x} \rangle - \sup_{i \in I} \{ \langle \boldsymbol{x}_i^*, \boldsymbol{x} \rangle - \alpha_i \} \right\} \\ &= \inf_{i \in I} \sup_{\boldsymbol{x} \in \mathscr{X}} \left\{ \langle \boldsymbol{x}^*, \boldsymbol{x} \rangle - \langle \boldsymbol{x}_i^*, \boldsymbol{x} \rangle + \alpha_i \right\}. \end{aligned}$$

对于 f^* 和 f^{**}，利用(2.2.10)可知，对任意的 $\boldsymbol{x} \in \mathscr{X}$ 和任意的 $i \in I$ 有

$$\begin{aligned} f^{**}(\boldsymbol{x}) &\geqslant \langle \boldsymbol{x}^*, \boldsymbol{x} \rangle - f^*(\boldsymbol{x}^*) \\ &= \langle \boldsymbol{x}^*, \boldsymbol{x} \rangle - \inf_{i \in I} \sup_{\boldsymbol{x} \in \mathscr{X}} \left\{ \langle \boldsymbol{x}^*, \boldsymbol{x} \rangle - \langle \boldsymbol{x}_i^*, \boldsymbol{x} \rangle + \alpha_i \right\} \\ &\geqslant \langle \boldsymbol{x}^*, \boldsymbol{x} \rangle - \sup_{\boldsymbol{y} \in \mathscr{X}} \left\{ \langle \boldsymbol{x}^*, \boldsymbol{y} \rangle - \langle \boldsymbol{x}_i^*, \boldsymbol{y} \rangle + \alpha_i \right\}. \end{aligned}$$

在最后的式子中取 $\boldsymbol{x}^* = \boldsymbol{x}_i^*$，则有

$$f^{**}(\boldsymbol{x}) \geqslant \langle \boldsymbol{x}^*, \boldsymbol{x} \rangle - a_i \quad \forall i \in I.$$

再根据(2.2.11)，得到

$$f^{**}(\boldsymbol{x}) \geqslant \sup_{i \in I}\{\langle \boldsymbol{x}_i^*, \boldsymbol{x} \rangle - a_i\} = f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \mathscr{X}.$$

由已证明的(1)，即得 $f^{**} = f$.

(3) 由(1)知 f^{**} 是 f 的弱函数，再据定理 2.2.11，得 f^* 是 $(f^{**})^*$ 的弱函数，即有

$$f^*(\boldsymbol{x}^*) \leqslant f^{***}(\boldsymbol{x}^*) \quad \forall \boldsymbol{x}^* \in \mathscr{X}^*. \qquad (2.2.12)$$

另外，在定理 2.2.12 的(1)中用 f^* 替换 f，则有

$$f^{***}(\boldsymbol{x}^*) \leqslant f^*(\boldsymbol{x}^*) \quad \forall \boldsymbol{x}^* \in \mathscr{X}^*. \qquad (2.2.13)$$

由(2.2.12)和(2.2.13)，便可得到结论. ▯

定理 2.2.13 设 $f: \mathscr{X} \to R \cup \{+\infty\}$ 是凸函数，$\boldsymbol{x} \in \mathscr{X}$，$\boldsymbol{x}^* \in \mathscr{X}^*$.

(1) 若 $a \in R$，$g(\boldsymbol{x}) = f(\boldsymbol{x}) + a$，则 $g^*(\boldsymbol{x}^*) = f^*(\boldsymbol{x}^*) - a$.

(2) 若 $a > 0$，$g(\boldsymbol{x}) = af(\boldsymbol{x})$，则 $g^*(\boldsymbol{x}^*) = af^*\left(\dfrac{\boldsymbol{x}^*}{a}\right)$.

(3) 若 $\boldsymbol{y} \in \mathscr{X}$，$g(\boldsymbol{x}) = f(\boldsymbol{x} + \boldsymbol{y})$，则 $g^*(\boldsymbol{x}^*) = f^*(\boldsymbol{x}^*) - \langle \boldsymbol{x}^*, \boldsymbol{y} \rangle$.

(4) 若 $\alpha \neq 0$，$g(\boldsymbol{x}) = f(\alpha \boldsymbol{x})$，则 $g^*(\boldsymbol{x}^*) = \dfrac{1}{\alpha} f^*(\boldsymbol{x}^*)$.

证明 (1) 由定义 2.2.4 的(1)，对任意的 $\boldsymbol{x}^* \in \mathscr{X}^*$ 有

$$\begin{aligned} g^*(\boldsymbol{x}^*) &= \sup_{\boldsymbol{x} \in \mathscr{X}}\{\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle - (f(\boldsymbol{x}) + a)\} \\ &= \sup_{\boldsymbol{x} \in \mathscr{X}}\{\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle - f(\boldsymbol{x})\} - a \\ &= f^*(\boldsymbol{x}^*) - a. \end{aligned}$$

(2) 根据定义 2.2.4 的(1)，直接有

$$g^*(\boldsymbol{x}^*) = \sup_{\boldsymbol{x} \in \mathscr{X}}\{\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle - af(\boldsymbol{x})\}$$

$$= a \sup_{x \in \mathscr{X}} \left\{ \left\langle \frac{x^*}{a}, x \right\rangle - f(x) \right\}$$

$$= af^* \left(\frac{x^*}{a} \right).$$

(3) 按定义 2.2.4 的(1),令 $z = x + y$,则得

$$g^*(x^*) = \sup_{x \in \mathscr{X}} \{ \langle x^*, x \rangle - f(x + y) \}$$

$$= \sup_{z \in \mathscr{X}} \{ \langle x^*, z - y \rangle - f(z) \}$$

$$= \sup_{z \in \mathscr{X}} \{ \langle x^*, z \rangle - f(z) \} - \langle x^*, y \rangle$$

$$= f^*(x^*) - \langle x^*, y \rangle.$$

(4) 由定义 2.2.4 的(1),令 $z = ax$,即可得证. □

定理 2.2.14　设 $f_i: \mathscr{X} \to R \cup \{+\infty\}$ ($i \in I$, I 是指标集)是凸函数,$\mathscr{X}^*$ 是 $\mathscr{X}$ 的对偶空间.

(1) 对任意的 $x^* \in \mathscr{X}^*$,有$\left(\inf_{i \in I} f_i \right)^*(x^*) = \sup_{i \in I} f_i^*(x^*)$.

(2) 对任意的 $x^* \in \mathscr{X}^*$,有$\left(\sup_{i \in I} f_i \right)^*(x^*) \leqslant \inf_{i \in I} f_i^*(x^*)$.

证明　由定义 2.2.4 的(1),对任意的 $x^* \in \mathscr{X}^*$ 有

$$\left(\inf_{i \in I} f_i \right)^*(x^*) = \sup_{x \in \mathscr{X}} \{ \langle x^*, x \rangle - \inf_{i \in I} f_i(x) \}$$

$$= \sup_{x \in \mathscr{X}} \left\{ \sup_{i \in I} \left[\langle x^*, x \rangle - f_i(x) \right] \right\}$$

$$= \sup_{i \in I} \left[\sup_{x \in \mathscr{X}} \{ \langle x^*, x \rangle - f_i(x) \} \right]$$

$$= \sup_{i \in I} f_i^*(x^*).$$

(2) 由定义 2.2.4 的(1),有

$$\left(\sup_{i \in I} f_i \right)^*(x^*) = \sup_{x \in \mathscr{X}} \{ \langle x^*, x \rangle - \sup_{i \in I} f_i(x) \}$$

$$= \sup_{x \in \mathscr{X}} \left\{ \inf_{i \in I} \left[\langle x^*, x \rangle - f_i(x) \right] \right\}$$

$$\leqslant \inf_{i \in I} \left[\sup_{x \in \mathscr{X}} \{ \langle x^*, x \rangle - f_i(x) \} \right]$$

$$= \inf_{i \in I} f_i^*(x^*). \square$$

最后，介绍一个基本的对偶定理. 为此，要引进凹函数的共轭函数的概念.

定义 2.2.5 设 $g: \mathscr{X} \to R \cup \{\pm \infty\}$ 是凹函数，$\mathscr{X}^*$ 是 $\mathscr{X}$ 的对偶空间. 令

$$g^*(x^*) = \inf_{x \in \mathscr{X}} \{ \langle x^*, x \rangle - g(x) \},$$

则称 $g^*: \mathscr{X}^* \to R \cup \{\pm \infty\}$ 是 g 的共轭函数.

与注 2.2.4 同样的理由，只需考虑 $g: \mathscr{X} \to R \cup \{-\infty\}$ 并且 $g \neq -\infty$ 的情形.

注 2.2.5 由定义 2.2.5 可以推知，对于凹函数的共轭函数，有与(2.2.10)类似的如下不等式：

$$g^*(x^*) + g(x) \leqslant \langle x^*, x \rangle \quad \forall x^* \in \mathscr{X}^*, \forall x \in \mathscr{X}. \tag{2.2.14}$$

关于凸函数和凹函数的共轭函数之间，有下面的对偶关系.

定理 2.2.15 设 $f: \mathscr{X} \to R \cup \{+\infty\}$ 是正常凸函数，$g: \mathscr{X} \to R \cup \{-\infty\}$ 是正常凹函数. 若 f 或 g 在某点 $x^0 \in \mathrm{dom} f \cap \mathrm{dom} g$ 处是连续的，则

$$\inf_{x \in \mathscr{X}} \{ f(x) - g(x) \} = \sup_{x^* \in \mathscr{X}^*} \{ g^*(x^*) - f^*(x^*) \}.$$

证明 由(2.2.14)和(2.2.10)，可知

$$g^*(x^*) + g(x) \leqslant \langle x^*, x \rangle \leqslant f^*(x^*) + f(x)$$

$$\forall x^* \in \mathscr{X}^*, \forall x \in \mathscr{X}.$$

据此，有

$$f(\boldsymbol{x}) - g(\boldsymbol{x}) \geqslant g^*(\boldsymbol{x}^*) - f^*(\boldsymbol{x}^*)$$

$$\forall \boldsymbol{x}^* \in \mathscr{X}^*, \forall \boldsymbol{x} \in \mathscr{X},$$

从而得

$$\inf_{\boldsymbol{x} \in \mathscr{X}} \{f(\boldsymbol{x}) - g(\boldsymbol{x})\} \geqslant \sup_{\boldsymbol{x}^* \in \mathscr{X}^*} \{g^*(\boldsymbol{x}^*) - f^*(\boldsymbol{x}^*)\}. \tag{2.2.15}$$

我们记

$$m = \inf_{\boldsymbol{x} \in \mathscr{X}} \{f(\boldsymbol{x}) - g(\boldsymbol{x})\}. \tag{2.2.16}$$

若 $m = -\infty$，则由(2.2.15)，定理的结论成立. 设 $m > -\infty$，作集合

$$S_1 = \{(\boldsymbol{x}, \eta) \in \mathscr{X} \times R \mid g(\boldsymbol{x}) + m \geqslant \eta\}, \tag{2.2.17}$$

$$S_2 = \mathrm{epi} f = \{(\boldsymbol{x}, \eta) \in \mathscr{X} \times R \mid \eta \geqslant f(\boldsymbol{x})\}. \tag{2.2.18}$$

从 $f \not\equiv +\infty$ 是凸函数，由定理 2.1.5 知 S_2 是凸集，由 $g \not\equiv -\infty$ 是凹函数，同理可以得知 S_1 是凸集. 根据定理的条件，不妨设 f 在某点 $\boldsymbol{x}^0 \in \mathrm{dom} f$ 处是连续的，则由 $f(\boldsymbol{x}^0) < +\infty$ 知，存在 η_0 使

$$\eta_0 = f(\boldsymbol{x}^0), \tag{2.2.19}$$

故由(2.2.18)有 $(\boldsymbol{x}^0, \eta_0) \in S_2$. 因为 f 在点 $\boldsymbol{x}^0$ 处是连续的，由定义 2.2.1 知存在 $\boldsymbol{x}^0$ 的邻域 $U(\boldsymbol{x}^0)$，对任意的 $\varepsilon > 0$，有

$$\left| f(\boldsymbol{x}) - f(\boldsymbol{x}^0) \right| < \varepsilon \quad \forall \boldsymbol{x} \in U(\boldsymbol{x}^0).$$

注意到(2.2.19)得

$$\eta_0 - \varepsilon < f(x) < \eta_0 + \varepsilon \quad \forall \boldsymbol{x} \in U(\boldsymbol{x}^0),$$

于是存在某点 $(\boldsymbol{x}', \eta') \in U(\boldsymbol{x}^0) \times (\eta_0 - \varepsilon, \eta_0 + \varepsilon)$. 据此，我们有 $(\boldsymbol{x}', \eta') \in \mathrm{int} S_2$，从而得 $\mathrm{int} S_2 \neq \varnothing$. 另外，由 $\boldsymbol{x}^0 \in \mathrm{dom} g$ 知 $g(\boldsymbol{x}^0) > -\infty$，因而存在 $\eta'' = g(\boldsymbol{x}^0) + m$. 由(2.2.17)知 $(\boldsymbol{x}^0, \eta'') \in S_1$，故得 $S_1 \neq \varnothing$. 现在再证明 $S_1 \cap \mathrm{int} S_2 = \varnothing$. 事实上若不然，则存在

$(\bar{x}, \bar{\eta}) \in S_1 \cap \mathrm{int}S_2$，于是 $(\bar{x}, \bar{\eta}) \in S_1$，从而由(2.2.17)知

$$\bar{\eta} \leqslant g(\bar{x}) + m. \tag{2.2.20}$$

又由 $(\bar{x}, \bar{\eta}) \in \mathrm{int}S_2$ 知存在$\bar{x}$的邻域$U(\bar{x})$和$\delta > 0$，使$U(\bar{x}) \times (\bar{\eta} - \delta, \bar{\eta} + \delta) \subset S_2$. 由此，按(2.2.18)知 $f(\bar{x}) \leqslant \bar{\eta} - \delta$，再由(2.2.20)得

$$f(\bar{x}) \leqslant \bar{\eta} - \delta < \bar{\eta} \leqslant g(\bar{x}) + m.$$

这说明存在 $\bar{x} \in \mathscr{X}$ 有 $m > f(\bar{x}) - g(\bar{x})$，导致与(2.2.16)相矛盾. 从以上得到的 $S_1 \neq \varnothing$，$\mathrm{int}S_2 \neq \varnothing$ 和 $S_1 \cap \mathrm{int}S_2 \neq \varnothing$，根据定理 1.3.7 的(1)，得知存在闭超平面正常分离 S_1 和 S_2，按定义 1.3.2 的(1)也即存在 $(x^*, \eta^*) \in (\mathscr{X}^* \times R)\setminus\{\theta\}$ 和 $\beta \in R$ 有

$$\langle (x^*, \eta^*), (x^1, \eta_1)\rangle \leqslant \beta \leqslant \langle (x^*, \eta^*), (x^2, \eta_2)\rangle$$

$$\forall (x^1, \eta_1) \in S_1,\ \forall (x^2, \eta_2) \in S_2,$$

即

$$\langle x^*, x^1\rangle + \langle \eta^*, \eta_1\rangle \leqslant \beta \leqslant \langle x^*, x^2\rangle + \langle \eta^*, \eta_2\rangle$$

$$\forall (x^1, \eta_1) \in S_1,\ \forall (x^2, \eta_2) \in S_2.$$

由(2.2.17)和(2.2.18)可取 $x = x^1 = x^2 \in \mathscr{X}$ 使 $\eta = f(x) = g(x) + m$，则得

$$\langle x^*, x\rangle + \eta^*[g(x) + m] \leqslant \beta \leqslant \langle x^*, x\rangle + \eta^* f(x)$$

$$\forall x \in \mathscr{X}.$$

因为 $\eta^* \neq 0$，记 $r = -\dfrac{\beta}{\eta^*} \in R$，$y^* = -\dfrac{x^*}{\eta^*} \in \mathscr{X}^*$，从上式推知

$$\langle y^*, x\rangle - f(x) \leqslant r,\ m + r \leqslant \langle y^*, x\rangle - g(x) \quad \forall x \in \mathscr{X},$$

于是由定义 2.2.4 的(1)和定义 2.2.5，有

$$f^*(y^*) = \sup_{x \in \mathscr{X}}\left[\langle y^*, x\rangle - f(x)\right] \leqslant r,$$

$$m + r \leqslant g^*(y^*) = \inf_{x \in \mathscr{X}} \left[\langle y^*, x \rangle - g(x) \right].$$

据此，我们得知 $m \leqslant g^*(y^*) - f^*(x^*)$，再由(2.2.15)和(2.2.16)有

$$m \leqslant g^*(y^*) - f^*(x^*) \leqslant \sup_{x^* \in \mathscr{X}^*} \{g^*(x^*) - f^*(x^*)\} \leqslant m,$$

因而

$$m = \sup_{x^* \in \mathscr{X}^*} \{g^*(x^*) - f^*(x^*)\}.$$

最后，注意到(2.2.16)，定理得证. ▯

考虑线性拓扑空间 $\mathscr{X}$ 中凸集 S 上的指示函数

$$\delta_S(x) = \begin{cases} 0, & x \in S; \\ +\infty, & x \in \mathscr{X} \setminus S. \end{cases}$$

在例 2.1.1 中，我们已指出它是凸函数. 从定义 2.2.4 的(1)可知，有

$$\delta_S^*(x^*) = \sup_{x \in \mathscr{X}} \{\langle x^*, x \rangle - \delta_S(x)\} = \sup_{x \in S} \langle x^*, x \rangle,$$

即

$$\delta_S^*(x^*) = \sup_{x \in S} \langle x^*, x \rangle, \ x^* \in \mathscr{X}^*. \tag{2.2.21}$$

据此，由定义 1.3.4 得知 δ_S 的共轭函数 $\delta_S^*: \mathscr{X}^* \to R \cup \{+\infty\}$ 是 S 的支撑函数. 由定理 2.2.10 的(1)可知，δ_S^* 总是下半连续的正常凸函数.

定理 2.2.16　设 $S \subset \mathscr{X}$ 是闭凸集，$\mathrm{int} S \neq \varnothing$，则

$$S = \{x \in \mathscr{X} \mid \langle x^*, x \rangle \leqslant \delta_S^*(x^*) \quad \forall x^* \in \mathscr{X}^*\}. \tag{2.2.22}$$

证明　因为 δ_S^* 是 S 的支撑函数. 由定理 1.3.13，注意到 S 是闭凸的即可得证. ▯

定理 2.2.16 表明任一闭凸集可以由其上的指示函数的对偶函数来表示.

§2.3 广义凸函数

这一节,阐述广义意义下的凸函数及其有关性质.我们主要介绍两类重要的广义凸函数——拟凸函数类和伪凸函数类.

先介绍线性空间上的拟凸函数类.

设 $\mathscr{X}$ 是线性空间.

定义 2.3.1 设 $S\subset\mathscr{X}$ 是非空凸集,$f:S\to R$ 是实值函数.

(1) 若对任意的 $\lambda\in(0,1)$ 有

$$f(\lambda\boldsymbol{x}^1+(1-\lambda)\boldsymbol{x}^2)\leqslant\max\{f(\boldsymbol{x}^1),f(\boldsymbol{x}^2)\}\quad\forall\boldsymbol{x}^1,\boldsymbol{x}^2\in S,\tag{2.3.1}$$

则称 f 是集合 S 上的拟凸函数,或函数 f 在 S 上是拟凸的.

(2) 若对任意的 $\lambda\in(0,1)$ 有

$$\begin{gathered}f(\lambda\boldsymbol{x}^1+(1-\lambda)\boldsymbol{x}^2)<\max\{f(\boldsymbol{x}^1),f(\boldsymbol{x}^2)\}\\ \forall\boldsymbol{x}^1,\boldsymbol{x}^2\in S,\ f(\boldsymbol{x}^1)\neq f(\boldsymbol{x}^2),\end{gathered}\tag{2.3.2}$$

则称 f 是集合 S 上的严格拟凸函数,或函数 f 在 S 上是严格拟凸的.

(3) 若对任意的 $\lambda\in(0,1)$ 有

$$\begin{gathered}f(\lambda\boldsymbol{x}^1+(1-\lambda)\boldsymbol{x}^2)<\max\{f(\boldsymbol{x}^1),f(\boldsymbol{x}^2)\}\\ \forall\boldsymbol{x}^1,\boldsymbol{x}^2\in S,\ \boldsymbol{x}^1\neq\boldsymbol{x}^2.\end{gathered}\tag{2.3.3}$$

则称 f 是集合 S 上的强拟凸函数,或函数 f 在 S 上是强拟凸的.

若 $-f$ 是 S 上的(严格,强)拟凸函数,则称 f 是 S 上的(严格,强)拟凹函数,或 f 在 S 上是(严格,强)拟凹的.

图 2.3.1 的(a)、(b)和(c)依次是单变量的拟凸函数、严格拟凸函数和强拟凸函数的图象.

注 2.3.1 在定义 2.3.1 中固定 $\boldsymbol{x}^1=\boldsymbol{x}^0\in S$,令 $\boldsymbol{x}^2=\boldsymbol{x}$,则(2.3.1)成为

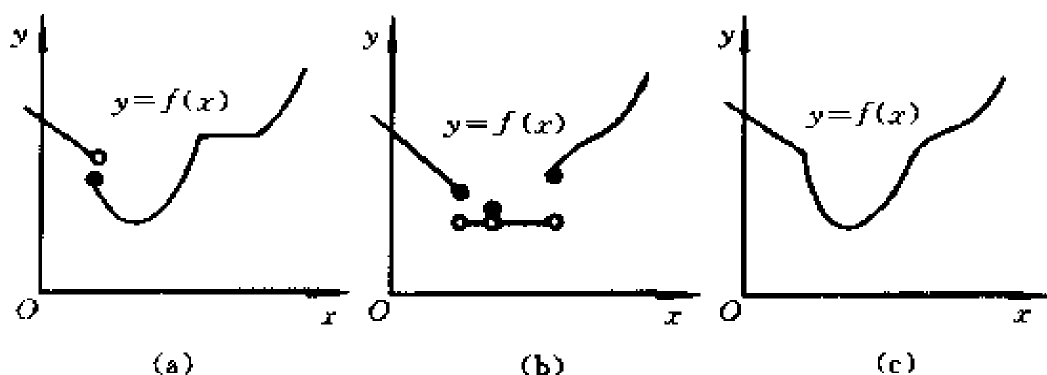


图　2.3.1

$$f(\lambda \boldsymbol{x}^0 + (1-\lambda)\boldsymbol{x}) \leqslant \max\{f(\boldsymbol{x}^0), f(\boldsymbol{x})\} \quad \forall \boldsymbol{x} \in S,$$

这时，称 f 在点 $\boldsymbol{x}^0$ 处相对于集合 S 是拟凸的. 同样地，若 $\boldsymbol{x}^1 = \boldsymbol{x}^0$ 和 $\boldsymbol{x}^2 = \boldsymbol{x}$ 分别满足相应的(2.3.2)和(2.3.3)，则称 f 在点 $\boldsymbol{x}^0$ 处相对于 S 是严格拟凸的和强拟凸的.

注 2.3.2　由定义 2.1.1 和定义 2.3.1 直接可以得知凸集上的凸函数是拟凸和严格拟凸的，严格凸函数是严格拟凸的，强凸函数是强拟凸的，反之，则不然. 从定义 2.3.1 还容易推知，强拟凸函数也是拟凸的和严格拟凸的. 然而，严格拟凸函数却不一定是拟凸的.

以下给出有关拟凸函数的两个基本结果.

定理 2.3.1　设 $S \subset \mathscr{R}^n$ 是非空凸集，则 $f: S \to R$ 是拟凸函数当且仅当对任何正整数 $m \geqslant 2$ 和任意的 $\lambda_i \geqslant 0\ (i = 1, \cdots, m)$，$\sum\limits_{i=1}^{m} \lambda_i = 1$，有

$$f\left(\sum_{i=1}^{m} \lambda_i \boldsymbol{x}^i\right) \leqslant \max\{f(\boldsymbol{x}^1), \cdots, f(\boldsymbol{x}^m)\} \quad \forall \boldsymbol{x}^1, \cdots, \boldsymbol{x}^m \in S.$$

证明　与定理 2.1.1 的证明类似，利用定义 2.3.1 的(1)可以推证. □

定理 2.3.2　设 $S \subset \mathscr{R}^n$ 是非空凸集，则 $f: S \to R$ 是拟凸函数当且仅当对任意的 $c \in R$，水平集 $H_S(f, c) = \{\boldsymbol{x} \in S \mid f(\boldsymbol{x}) \leqslant c\}$ 是

凸集.

证明 必要性. 任取 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in H_S(f, c) \subset S$, 则有 $f(\boldsymbol{x}^1) \leqslant c$ 和 $f(\boldsymbol{x}^2) \leqslant c$. 因为 S 是凸集, 故对任意的 $\lambda \in (0, 1)$ 有 $\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 \in S$. 由于 f 是 S 上的拟凸函数, 按定义 2.3.1 的(1)得

$$f(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \leqslant \max\{f(\boldsymbol{x}^1), f(\boldsymbol{x}^2)\} \leqslant c,$$

于是 $\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 \in H_S(f, c)$. 据此, 由定义 1.1.1 可知 $H_S(f, c)$ 是凸集.

充分性. 对任意的 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in S$, 取

$$c = \max\{f(\boldsymbol{x}^1), f(\boldsymbol{x}^2)\}, \tag{2.3.4}$$

则 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in H_S(f, c)$. 因为 $H_S(f, c)$ 是凸集, 故对任意的 $\lambda \in (0, 1)$, 有 $\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 \in H_S(f, c)$, 于是得 $f(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \leqslant c$. 据此, 由(2.3.4)按定义 2.3.1 的(1), 便知 f 在 S 上是拟凸的. ▯

注 2.3.3 由定理 2.1.2 已知, 函数的相应任意水平集为凸集是该函数为凸函数的必要条件. 定理 2.3.2 则进一步表明了函数的相应任意水平集为凸集是该函数为拟凸的充要条件. 从而可见, 函数的拟凸性更充分地反映了其相应水平集的凸性特征.

下面的定理指出, 在线性拓扑空间上为下半连续的严格拟凸函数必定是拟凸的.

设 $\mathscr{X}$ 是线性拓扑空间.

定理 2.3.3 设 $S \subset \mathscr{X}$ 是非空凸集. 若 $f: S \to R$ 是下半连续的严格拟凸函数, 则 f 是 S 上的拟凸函数.

证明 设对任意的 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in S$, 有 $f(\boldsymbol{x}^1) \neq f(\boldsymbol{x}^2)$. 由 f 在 S 上是严格拟凸的, 从(2.3.2)可得(2.3.1). 因此, 根据定义 2.3.1 中的(1)即知 f 在 S 上是拟凸的.

设存在 $\bar{\boldsymbol{x}}^1, \bar{\boldsymbol{x}}^2 \in S$ 有 $f(\bar{\boldsymbol{x}}^1) = f(\bar{\boldsymbol{x}}^2)$. 我们证明这时条件(2.3.1)成立, 即对任意的 $\lambda \in (0, 1)$ 有

$$f(\lambda \bar{\boldsymbol{x}}^1 + (1-\lambda)\bar{\boldsymbol{x}}^2) \leqslant f(\bar{\boldsymbol{x}}^1) = f(\bar{\boldsymbol{x}}^2).$$

为此，与定理 2.1.3 的证明类似，用反证法，假设对某个 $\lambda_0 \in (0, 1)$ 有

$$f(\lambda_0 \bar{x}^1 + (1-\lambda_0)\bar{x}^2) > f(\bar{x}^1) = f(\bar{x}^2).$$

因为 S 是凸集，故 $x^0 = \lambda_0 \bar{x}^1 + (1-\lambda_0)\bar{x}^2 \in S$. 由上式得

$$f(\bar{x}^1) < f(x^0), \quad f(\bar{x}^2) < f(x^0). \tag{2.3.5}$$

取充分小的 $\varepsilon_0 > 0$，使(2.3.5)的第一式有 $f(\bar{x}^1) + \varepsilon_0 < f(x^0)$，即

$$f(\bar{x}^1) < f(x^0) - \varepsilon_0.$$

因为 f 在 $x^0 \in S$ 处是下半连续的，由定义 2.1.2 的(1)知，对任意的 $\epsilon > 0$，存在 λ_0 的邻域 $U(\lambda_0)$，有

$$-\epsilon < f(\lambda \bar{x}^1 + (1-\lambda)\bar{x}^2) - f(x^0) \quad \forall \lambda \in U(\lambda_0).$$

据此，由以上两式得知

$$f(\bar{x}^1) < f(x^0) - \varepsilon_0 < f(\lambda \bar{x}^1 + (1-\lambda)\bar{x}^2) \quad \forall \lambda \in U(\lambda_0).$$

取 $\lambda' \in U(\lambda_0)$，$\lambda' \neq \lambda_0$，记 $x' = \lambda' \bar{x}^1 + (1-\lambda')\bar{x}^2$，则有

$$f(\bar{x}^1) < f(x'). \tag{2.3.6}$$

不妨设 $x^0 \in (\bar{x}^1, x')$，则 $x' \in (x^0, \bar{x}^2)$. 因为 f 是严格拟凸的，所以由定义 2.3.1 的(2)分别注意到(2.3.6)和(2.3.5)的第 2 式得到

$$f(x^0) < \max\{f(\bar{x}^1), f(x')\} = f(x'),$$

$$f(x') < \max\{f(x^0), f(\bar{x}^2)\} = f(x^0),$$

但这两式是矛盾的. 定理得证. ▯

拟凸函数的以下一些性质与凸函数的相应性质相类似.

定理 2.3.4　设 $S \subset \mathscr{R}^n$ 是非空凸集，$f_i: S \to R$ 是拟凸函数($i \in I$，I 是指标集). 若

$$f(x) = \sup_{i \in I} f_i(x), \tag{2.3.7}$$

则 $f: S \to R$ 是拟凸函数.

证明 由 $f_i (i \in I)$ 是拟凸的，根据定理 2.3.2 知，对任意的 $c \in R$，水平集 $H_S(f_i, c) = \{x \in S \mid f_i(x) \leqslant c\}$ $(i \in I)$ 是凸的. 因此，由定理 1.1.2 的(2)得知 $\bigcap_{i \in I} H_S(f_i, c)$ 是凸集. 任取 $x \in \bigcap_{i \in I} H_S(f_i, c)$，则 $f_i(x) \leqslant c$ $(i \in I)$，由(2.3.7)，它意味着 $f(x) \leqslant c$，故得

$$H_S(f, c) = \{x \in S \mid f(x) \leqslant c\} = \bigcap_{i \in I} H_S(f_i, c).$$

由于对任意的 $c \in R$ 水平集 $H_S(f, c)$ 是凸的，再根据定理 2.3.2 即得 f 是 S 上的拟凸函数. □

定理 2.3.5 设 $f: \mathscr{Y} \to R$ 和 $g: R \to R$ 是实值函数，并且 $g(t) = f(x + td)$（其中 $d \in \mathscr{Y}$），则 f 是拟凸函数当且仅当对任意的 $d \in \mathscr{Y}$，g 是拟凸函数.

证明 利用定义 2.3.1 的(1)，与定理 2.1.8 的证明类似可得证. □

定理 2.3.6 设 $h: \mathscr{Y} \to R$ 是拟凸函数，$g: R \to R$ 是非减函数，$f(x) = g(h(x))$，则 $f: \mathscr{Y} \to R$ 是拟凸函数.

证明 任取 $x^1, x^2 \in \mathscr{Y}$ 和 $\lambda \in (0, 1)$，因为 h 是拟凸的和 g 是非减的，有

$$\begin{aligned} f(\lambda x^1 + (1-\lambda)x^2) &= g(h(\lambda x^1 + (1-\lambda)x^2)) \\ &\leqslant g\left(\max\{h(x^1), h(x^2)\}\right) \\ &= \max\{g(h(x^1)), g(h(x^2))\} \\ &= \max\{f(x^1), f(x^2)\}. \end{aligned}$$

由此，根据定义 2.3.1 的(1)得知 f 是拟凸的. □

注 2.3.4 定理 2.3.6 与定理 2.1.9 相似. 但要注意，在定理 2.3.6 中并不要求 g 具有凸性或拟凸性.

再引进另一类由广义导数定义的广义凸函数的概念. 为此，先给出映射的 Gâteaux 导数的定义.

再设 $\mathscr{Y}$ 是线性拓扑空间.

定义 2.3.2 设 $\varphi: \mathscr{X} \to \mathscr{Y}$ 是映射,点 $\boldsymbol{x}^0 \in \mathscr{X}$. 若对任意的 $\boldsymbol{d} \in \mathscr{X}$,极限

$$\varphi'_{x^0}(\boldsymbol{d}) = \lim_{t \to 0} \frac{\varphi(\boldsymbol{x}^0 + t\boldsymbol{d}) - \varphi(\boldsymbol{x}^0)}{t} \tag{2.3.8}$$

存在,则称 φ 在点 $\boldsymbol{x}^0$ 处是 Gâteaux 可微的. 映射 $\varphi'_{x^0}: \mathscr{X} \to \mathscr{Y}$, $\boldsymbol{d} \mapsto \varphi'_{x^0}(\boldsymbol{d})$ 称为是 φ 在点 $\boldsymbol{x}^0$ 处的 Gâteaux 导数. 若 φ 在集合 $S \subset \mathscr{X}$ 的每一点处都是 Gâteaux 可微的,则称 φ 在 S 上是 Gâteaux 可微的.

特别地,若 $\varphi: \mathscr{X} \to R \cup \{\pm\infty\}$ 是广义实值函数,点 $\boldsymbol{x}^0 \in \text{int}(\text{dom}\varphi)$,则 φ 在点 $\boldsymbol{x}^0$ 处的 Gâteaux 导数 $\varphi'_{x^0}: \mathscr{X} \to R$ 是实值函数.

注 2.3.5 由定义 2.3.2 易知,对于 $\boldsymbol{x} \in \mathscr{X}$,有

$$\varphi'_x(\boldsymbol{0}) = \boldsymbol{0}, \tag{2.3.9}$$

$$\varphi'_x(\alpha\boldsymbol{d}) = \alpha\varphi'_x(\boldsymbol{d}), \ \alpha \in R. \tag{2.3.10}$$

定理 2.3.7 设 $f: \mathscr{X} \to R \cup \{\pm\infty\}$ 是广义实值函数,点 $\boldsymbol{x}^0 \in \text{int}(\text{dom}f)$. 若 f 在点 $\boldsymbol{x}^0$ 处是 Gâteaux 可微的,则 f 在点 $\boldsymbol{x}^0$ 处是连续的.

证明 由 f 在点 $\boldsymbol{x}^0$ 处是 Gâteaux 可微的按定义 2.3.2 可知,对任意的 $\boldsymbol{d} \in \mathscr{X}$,极限

$$f'_{x^0}(\boldsymbol{d}) = \lim_{t \to 0} \frac{f(\boldsymbol{x}^0 + t\boldsymbol{d}) - f(\boldsymbol{x}^0)}{t}$$

存在. 据此,我们有

$$f(\boldsymbol{x}^0 + t\boldsymbol{d}) - f(\boldsymbol{x}^0) = f'_{x^0}(\boldsymbol{d})t + o(t),$$

其中 $o(t) \to 0 \ (t \to 0)$. 记 $\boldsymbol{x} = \boldsymbol{x}^0 + t\boldsymbol{d}$,则对任意的 $\varepsilon > 0$,只要 t 充分小便存在 $\boldsymbol{x}^0$ 的邻域 $U(\boldsymbol{x}^0)$,使

$$|f(\boldsymbol{x}) - f(\boldsymbol{x}^0)| = |f'_{x^0}(\boldsymbol{d})t + o(t)| < \varepsilon \quad \forall \boldsymbol{x} \in U(\boldsymbol{x}^0).$$

于是，由定义 2.2.1，即知 f 在点 $\boldsymbol{x}^0$ 处是连续的. ▯

以下介绍线性拓扑空间上的伪凸函数类.

定义 2.3.3 设 $S\subset\mathscr{X}$ 是非空开凸集，实值函数 $f: S\to R$ 是 Gâteaux 可微的.

(1) 若有

$$f(\boldsymbol{x}^1)>f(\boldsymbol{x}^2)\Rightarrow f'_{\boldsymbol{x}^1}(\boldsymbol{x}^2-\boldsymbol{x}^1)<0\quad \forall\boldsymbol{x}^1,\boldsymbol{x}^2\in S, \tag{2.3.11}$$

则称 f 是集合 S 上的伪凸函数，或函数 f 在 S 上是伪凸的.

(2) 若有

$$f(\boldsymbol{x}^1)\geqslant f(\boldsymbol{x}^2)\Rightarrow f'_{\boldsymbol{x}^1}(\boldsymbol{x}^2-\boldsymbol{x}^1)<0\quad \forall\boldsymbol{x}^1,\boldsymbol{x}^2\in S,\ \boldsymbol{x}^1\neq\boldsymbol{x}^2, \tag{2.3.12}$$

则称 f 是集合 S 上的强伪凸函数，或函数 f 在 S 上是强伪凸的.

(3) 若有

$$f(\boldsymbol{x}^1)\geqslant f(\boldsymbol{x}^2)\Rightarrow f'_{\boldsymbol{x}^1}(\boldsymbol{x}^2-\boldsymbol{x}^1)\leqslant 0\quad \forall\boldsymbol{x}^1,\boldsymbol{x}^2\in S, \tag{2.3.13}$$

则称 f 是集合 S 上的次伪凸函数，或函数 f 在 S 上是次伪凸的.

若 $-f$ 是 S 上的(强，次)伪凸函数，则称 f 是 S 上的(强，次)伪凹函数，或 f 在 S 上是(强，次)伪凹的.

图 2.3.2 的(a)、(b)和(c)依次是单变量的伪凸函数、强伪凸

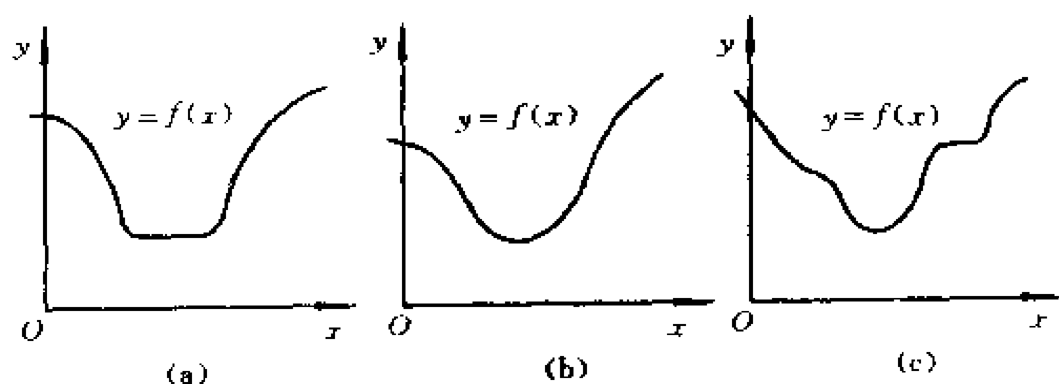


图 2.3.2

函数和次伪凸函数的图象.

注 2.3.6　在定义 2.3.3 中固定 $x^1 = x^0 \in S$，令 $x^1 = x$，则对应于定义 2.3.3 的(1)、(2)和(3)，依次称 f 在点 x^0 处相对于集合 S 是伪凸的、强伪凸的和次伪凸的.

注 2.3.7　由定义 2.3.3 不难得知，若 f 是 S 上的强伪凸(凹)函数，则 f 是 S 上的伪凸(凹)函数和次伪凸(凹)函数，但反之不然.

注 2.3.8　对于有限维 Euclid 空间 R^n 中开凸集 S 上的可微函数 $f: S \to R$，由定义 2.3.2 可知，f 在点 x^1 处的 Gâteaux 导数在 $d = x^2 - x^1$ 处的值 $f'_{x^1}(x^2 - x^1)$，即为 f 在点 x^1 处沿方向 $d = x^2 - x^1$ 的方向导数，故有 $f'_{x^1}(x^2 - x^1) = \nabla f(x^1)^T(x^2 - x^1)$ (其中 $\nabla f(x^1)$ 是 f 在点 x^1 处的梯度). 这时，f 在 S 上是伪凸的意味着

$$f(x^1) > f(x^2) \Rightarrow \nabla f(x^1)^T(x^2 - x^1) < 0 \quad \forall x^1, x^2 \in S,$$

f 在 S 上是强伪凸的意味着

$$f(x^1) \geqslant f(x^2) \Rightarrow \nabla f(x^1)^T(x^2 - x^1) < 0$$

$$\forall x^1, x^2 \in S, x^1 \neq x^2,$$

f 在 S 上是次伪凸的意即

$$f(x^1) \geqslant f(x^2) \Rightarrow \nabla f(x^1)^T(x^2 - x^1) \leqslant 0 \quad \forall x^1, x^2 \in S.$$

下面的定理指出，线性拓扑空间上 Gâteaux 可微的凸函数和强凸函数分别也是伪凸函数和强伪凸函数.

定理 2.3.8　设 $S \subset \mathscr{X}$ 是非空开凸集，$f: S \to R$ 是 Gâteaux 可微的.

(1) 若 f 是 S 上的凸函数，则 f 是 S 上的伪凸函数和次伪凸函数.

(2) 若 f 是 S 上的强凸函数，则 f 是 S 上的强伪凸函数.

证明　(1) 因 f 是 S 上的凸函数，按定义 2.1.1 的(1)，对于

任意的 $x^1, x^2 \in S$ 和任意的 $t \in (0, 1)$，有

$$\begin{aligned} f(x^1 + t(x^2 - x^1)) - f(x^1) &= f(tx^2 + (1-t)x^1) - f(x^1) \\ &\leqslant tf(x^2) + (1-t)f(x^1) - f(x^1) \\ &= t\left[f(x^2) - f(x^1)\right]. \end{aligned}$$

由此，按定义 2.3.2 得

$$\begin{aligned} f'_{x^1}(x^2 - x^1) &= \lim_{t \to 0} \frac{f(x^1 + t(x^2 - x^1)) - f(x^1)}{t} \\ &\leqslant f(x^2) - f(x^1). \end{aligned} \tag{2.3.14}$$

依据(2.3.14)，从 $f(x^1) > f(x^2)$ 即可推出 $f'_{x^1}(x^2 - x^1) < 0$，于是按定义 2.3.3 的(1)得知，f 是 S 上的伪凸函数.

又据(2.3.14)，从 $f(x^1) \geqslant f(x^2)$ 可以推出 $f'_{x^1}(x^2 - x^1) \leqslant 0$，故按定义 2.3.3 的(3)，即知 f 是 S 上的次伪凸函数.

(2) 与(1)的证明类似，由定义 2.3.2 和定义 2.1.1 的(3)可以推知，对任意的 $x^1, x^2 \in S$，$x^1 \neq x^2$，有

$$f'_{x^1}(x^2 - x^1) < f(x^2) - f(x^1).$$

由此，从 $f(x^1) \geqslant f(x^2)$ 可推出 $f'_{x^1}(x^2 - x^1) < 0$，故由定义 2.3.3 中的(2)，即得 f 是 S 上的强伪凸函数. ▯

以下讨论函数的伪凸性与拟凸性之间的几个关系.

定理 2.3.9 设 $S \subset \mathscr{X}$ 是非空开凸集，$f: S \to R$ 是 Gâteaux 可微的.

(1) 若 f 是 S 上的伪凸函数，则 f 是 S 上的严格拟凸函数.

(2) 若 f 是 S 上的伪凸函数，则 f 是 S 上的拟凸函数.

证明 (1) 用反证法，假设 f 在 S 上不是严格拟凸的，则由定义 2.3.1 的(2)得知，存在 $x^1, x^2 \in S$，$f(x^1) \neq f(x^2)$，对某 $\lambda_0 \in (0, 1)$ 有

$$f(\lambda_0 x^1 + (1 - \lambda_0)x^2) \geqslant \max\{f(x^1), f(x^2)\}. \tag{2.3.15}$$

不妨设

$$f(\boldsymbol{x}^1) > f(\boldsymbol{x}^2), \tag{2.3.16}$$

则从(2.3.15)知

$$f(\lambda_0\boldsymbol{x}^1 + (1-\lambda_0)\boldsymbol{x}^2) \geqslant f(\boldsymbol{x}^1) > f(\boldsymbol{x}^2). \tag{2.3.17}$$

由 S 是凸集，f 在闭线段 $[\boldsymbol{x}^1, \boldsymbol{x}^2] \subset S$ 上是 Gâteaux 可微的，根据定理 2.3.7，它在该线段上是连续的. 因此，f 在 $[\boldsymbol{x}^1, \boldsymbol{x}^2]$ 上有最大值. 从(2.3.17)和 $\lambda_0\boldsymbol{x}^1 + (1-\lambda_0)\boldsymbol{x}^2 \in (\boldsymbol{x}^1, \boldsymbol{x}^2)$ 可知，该最大值可在 $(\boldsymbol{x}^1, \boldsymbol{x}^2)$ 内取到，于是存在点 $\bar{\boldsymbol{x}} \in (\boldsymbol{x}^1, \boldsymbol{x}^2)$，对任意的 $t \in (0, 1)$ 有

$$f(\bar{\boldsymbol{x}}) \geqslant f\big(\bar{\boldsymbol{x}} + t(\boldsymbol{x}^1 - \bar{\boldsymbol{x}})\big),$$

$$f(\bar{\boldsymbol{x}}) \geqslant f\big(\bar{\boldsymbol{x}} + t(\boldsymbol{x}^2 - \bar{\boldsymbol{x}})\big).$$

据此，由(2.3.8)得

$$f'_{\bar{x}}(\boldsymbol{x}^1 - \bar{\boldsymbol{x}}) = \lim_{t\to 0}\frac{f\big(\bar{\boldsymbol{x}} + t(\boldsymbol{x}^1 - \bar{\boldsymbol{x}})\big) - f(\bar{\boldsymbol{x}})}{t} \leqslant 0, \tag{2.3.18}$$

$$f'_{\bar{x}}(\boldsymbol{x}^2 - \bar{\boldsymbol{x}}) = \lim_{t\to 0}\frac{f\big(\bar{\boldsymbol{x}} + t(\boldsymbol{x}^2 - \bar{\boldsymbol{x}})\big) - f(\bar{\boldsymbol{x}})}{t} \leqslant 0. \tag{2.3.19}$$

设 $\bar{\boldsymbol{x}} = \bar{\lambda}\boldsymbol{x}^1 + (1-\bar{\lambda})\boldsymbol{x}^2$ (其中 $\bar{\lambda} \in (0, 1)$)，则

$$\boldsymbol{x}^1 - \bar{\boldsymbol{x}} = -\frac{1-\bar{\lambda}}{\bar{\lambda}}(\boldsymbol{x}^2 - \bar{\boldsymbol{x}}).$$

于是，由(2.3.18)和(2.3.19)利用(2.3.10)有

$$0 \geqslant f'_{\bar{x}}(\boldsymbol{x}^1 - \bar{\boldsymbol{x}}) = f'_{\bar{x}}\left(-\frac{1-\bar{\lambda}}{\bar{\lambda}}(\boldsymbol{x}^2 - \bar{\boldsymbol{x}})\right)$$

$$= -\frac{1-\bar{\lambda}}{\bar{\lambda}}f'_{\bar{x}}(\boldsymbol{x}^2 - \bar{\boldsymbol{x}}) \geqslant 0,$$

从而得

$$f'_{\bar{x}}(x^2-\bar{x})=0. \tag{2.3.20}$$

因为已知 f 在 S 上是伪凸的，所以由(2.3.20)可推知 $f(\bar{x})\leqslant f(x^2)$(否则，若 $f(\bar{x})>f(x^2)$，由定义 2.3.3 的(1)得 $f'_{\bar{x}}(x^2-\bar{x})<0$，与(2.3.20)矛盾). 注意到(2.3.16)得 $f(x^1)>f(\bar{x})$，因为 f 是连续的，这与 $f(\bar{x})$ 是 $(x^1,\ x^2)$ 中的最大值相矛盾.

(2) 由 f 在 S 上是 Gâteaux 可微的，它在 S 上是连续的，从而也是下半连续的. 于是，由已证得的(1)和定理 2.3.3，即知 f 在 S 上是拟凸的. ▯

定理 2.3.10 设 $S\subset\mathscr{X}$ 是非空开凸集，$f\colon S\to R$ 是 Gâteaux 可微的. 若 f 是 S 上的强伪凸函数，则 f 是 S 上的强拟凸函数.

证明 用反证法. 假设 f 在 S 上不是强拟凸的，则由定义 2.3.1 的(3)得知，存在 $\bar{x}^1,\ \bar{x}^2\in S,\ \bar{x}^1\neq\bar{x}^2$，对某个 $\bar{\lambda}\in(0,\ 1)$ 有

$$f(\bar{\lambda}\bar{x}^1+(1-\bar{\lambda})\bar{x}^2)\geqslant\max\{f(\bar{x}^1),\ f(\bar{x}^2)\}. \tag{2.3.21}$$

设 $\bar{x}=\bar{\lambda}\bar{x}^1+(1-\bar{\lambda})\bar{x}^2\neq\bar{x}^1$，由(2.3.21)有 $f(\bar{x})\geqslant f(\bar{x}^1)$，于是根据 f 是强伪凸函数，按定义 2.3.3 的(2)得

$$f'_{\bar{x}}(\bar{x}^1-\bar{x})<0. \tag{2.3.22}$$

同理，从(2.3.21)有 $f(\bar{x})\geqslant f(\bar{x}^2)$，从而又得

$$f'_{\bar{x}}(\bar{x}^2-\bar{x})<0. \tag{2.3.23}$$

因为 $\bar{x}^1-\bar{x}=-\dfrac{1-\bar{\lambda}}{\bar{\lambda}}(\bar{x}^2-\bar{x})$，所以由(2.3.22)和(2.3.23)利用(2.3.10)推知

$$0>f'_{\bar{x}}(\bar{x}^1-\bar{x})=-\frac{1-\bar{\lambda}}{\bar{\lambda}}f'_{\bar{x}}(\bar{x}^2-\bar{x})>0,$$

导致矛盾. ▯

定理 2.3.11 设 $S\subset\mathscr{X}$ 是非空开凸集，$f\colon S\to R$ 是

Gâteaux 可微的，则 f 是 S 上的拟凸函数当且仅当 f 是 S 上的次伪凸函数.

证明　必要性. 由 f 在 S 上是拟凸的，按定义 2.3.1 的(1)知，对任意的 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in S$ 和任意的 $\lambda \in (0, 1)$ 有

$$f(\lambda \boldsymbol{x}^2 + (1-\lambda)\boldsymbol{x}^1) \leqslant \max\{f(\boldsymbol{x}^2), f(\boldsymbol{x}^1)\}.$$

设 $f(\boldsymbol{x}^1) \geqslant f(\boldsymbol{x}^2)$，从上式知

$$\begin{aligned} & f(\boldsymbol{x}^1 + \lambda(\boldsymbol{x}^2 - \boldsymbol{x}^1)) - f(\boldsymbol{x}^1) \\ = & f(\lambda \boldsymbol{x}^2 + (1-\lambda)\boldsymbol{x}^1) - f(\boldsymbol{x}^1) \leqslant 0. \end{aligned}$$

据此，由定义 2.3.2 得到

$$f'_{x^1}(\boldsymbol{x}^2 - \boldsymbol{x}^1) = \lim_{\lambda \to 0} \frac{f(\boldsymbol{x}^1 + \lambda(\boldsymbol{x}^2 - \boldsymbol{x}^1)) - f(\boldsymbol{x}^1)}{\lambda} \leqslant 0.$$

于是，按定义 2.3.3 的(3)即知 f 在 S 上是次伪凸的.

充分性. 已知 f 在 S 上是次伪凸的，用反证法，假设 f 在 S 上不是拟凸的. 从后者由定义 2.3.1 中的(1)可知，存在 $\bar{\boldsymbol{x}}^1, \bar{\boldsymbol{x}}^2 \in S$ 和 $\lambda_0 \in (0, 1)$ 有

$$f(\lambda_0 \bar{\boldsymbol{x}}^1 + (1-\lambda_0)\bar{\boldsymbol{x}}^2) > \max\{f(\bar{\boldsymbol{x}}^1), f(\bar{\boldsymbol{x}}^2)\}. \tag{2.3.24}$$

因为 S 是凸集，故

$$\boldsymbol{x}^0 = \lambda_0 \bar{\boldsymbol{x}}^1 + (1-\lambda_0)\bar{\boldsymbol{x}}^2 \in S. \tag{2.3.25}$$

现设 $f(\bar{\boldsymbol{x}}^1) \geqslant f(\bar{\boldsymbol{x}}^2)$，从(2.3.24)和(2.3.25)可得

$$f(\bar{\boldsymbol{x}}^2) \leqslant f(\bar{\boldsymbol{x}}^1) < f(\boldsymbol{x}^0). \tag{2.3.26}$$

由于 f 在 S 上是次伪凸的，根据(2.3.26)和定义 2.3.3 的(3)推知

$$f'_{x^0}(\bar{\boldsymbol{x}}^1 - \boldsymbol{x}^0) \leqslant 0,\ f'_{x^0}(\bar{\boldsymbol{x}}^2 - \boldsymbol{x}^0) \leqslant 0.$$

由(2.3.25),有 $\bar{\boldsymbol{x}}^1 - \boldsymbol{x}^0 = (1-\lambda_0)(\bar{\boldsymbol{x}}^1 - \bar{\boldsymbol{x}}^2)$ 和 $\bar{\boldsymbol{x}}^2 - \boldsymbol{x}^0 = \lambda_0(\bar{\boldsymbol{x}}^2 - \bar{\boldsymbol{x}}^1)$,代入以上两式,注意到(2.3.10)有 $f'_{x^0}(\bar{\boldsymbol{x}}^1 - \bar{\boldsymbol{x}}^2) \leqslant 0$ 和 $f'_{x^0}(\bar{\boldsymbol{x}}^2 - \bar{\boldsymbol{x}}^1) \leqslant 0$,从而得

$$f'_{x^0}(\bar{\boldsymbol{x}}^2 - \bar{\boldsymbol{x}}^1) = 0. \tag{2.3.27}$$

记

$$\lambda' = \inf\{\lambda \in [0,1] \mid f(\lambda\bar{\boldsymbol{x}}^1 + (1-\lambda)\boldsymbol{x}^0) \leqslant f(\bar{\boldsymbol{x}}^1)\}, \tag{2.3.28}$$

则有 $\lambda' > 0$. 事实上,若有 $\lambda' = 0$,则由 f 在 S 上是 Gâteaux 可微的,根据定理 2.3.7,它在 S 上是连续的,从而也是下半连续的.由定理 2.2.5 的(1)和(3)知,$D = \{\boldsymbol{x} \in S \mid f(\boldsymbol{x}) \leqslant f(\bar{\boldsymbol{x}}^1)\}$ 是闭集.据此,设 $\{\lambda_k\} \subset [0,1]$,$\lambda_k\bar{\boldsymbol{x}}^1 + (1-\lambda_k)\boldsymbol{x}^0 \in D$,并且 $\lambda_k \to \lambda' = 0$ $(k \to \infty)$,则 $\lambda_k\bar{\boldsymbol{x}}^1 + (1-\lambda_k)\boldsymbol{x}^0 \to \boldsymbol{x}^0 \in D$,即有 $f(\boldsymbol{x}^0) \leqslant f(\bar{\boldsymbol{x}}^1)$,它与(2.3.26)相矛盾.现令

$$\boldsymbol{x}' = \lambda'\bar{\boldsymbol{x}}^1 + (1-\lambda')\boldsymbol{x}^0, \tag{2.3.29}$$

则 $\boldsymbol{x}' \in (\bar{\boldsymbol{x}}^1, \boldsymbol{x}^0)$,并且由(2.3.28)有

$$f(\boldsymbol{x}') \leqslant f(\bar{\boldsymbol{x}}^1). \tag{2.3.30}$$

另外,由(2.3.28)知

$$f(\boldsymbol{x}) > f(\bar{\boldsymbol{x}}^1) \quad \forall \boldsymbol{x} \in (\boldsymbol{x}', \boldsymbol{x}^0). \tag{2.3.31}$$

从 f 在 S 上是 Gâteaux 可微的,根据中值定理[6],有

$$f(\boldsymbol{x}^0) - f(\boldsymbol{x}') = f'_{\tilde{x}}(\boldsymbol{x}^0 - \boldsymbol{x}'),$$

其中 $\tilde{\boldsymbol{x}} \in (\boldsymbol{x}', \boldsymbol{x}^0)$. 由(2.3.25)和(2.3.29)知 $\boldsymbol{x}^0 - \boldsymbol{x}' = (1-\lambda_0)\lambda'(\bar{\boldsymbol{x}}^2 - \bar{\boldsymbol{x}}^1)$. 代入上式并注意(2.3.10),得

$$f(\boldsymbol{x}^0) - f(\boldsymbol{x}') = (1-\lambda_0)\lambda' f'_{\tilde{x}}(\bar{\boldsymbol{x}}^2 - \bar{\boldsymbol{x}}^1). \tag{2.3.32}$$

从 $\tilde{\boldsymbol{x}} \in (\boldsymbol{x}', \boldsymbol{x}^0)$ 和(2.3.31)知 $f(\bar{\boldsymbol{x}}^1) < f(\tilde{\boldsymbol{x}})$,又 $\tilde{\boldsymbol{x}}$ 也是 $\bar{\boldsymbol{x}}^1$ 和 $\bar{\boldsymbol{x}}^2$ 的凸组合,因此对(2.3.26)至(2.3.27)以 $\tilde{\boldsymbol{x}}$ 替换 $\boldsymbol{x}^0$,可以推得

$$f'_{\bar{x}}(\bar{x}^2 - \bar{x}^1) = 0.$$

将它代入(2.3.32),再注意(2.3.30),便得 $f(x^0) = f(x') \leqslant f(\bar{x}^1)$,这导致与(2.3.26)矛盾. ▯

§ 2.4 凸映射和广义凸映射

现在将函数的凸性概念推广到映射的情况,论述凸映射和广义凸映射及其有关性质和关系.由于映射值空间上的序由该空间中的锥来确定,故论述映射的凸性要依赖于相应的序锥.因此,有关凸映射也称为锥凸映射.

设 $\mathscr{X}$ 是线性空间,$\mathscr{Y}$ 是序线性拓扑空间.除特别说明外,$K \subset \mathscr{Y}$ 是内部非空的尖闭凸锥,$\mathscr{Y}$ 中的序由 K 确定.

先介绍锥凸映射类.

定义 2.4.1 设 $S \subset \mathscr{X}$ 是非空凸集,$\varphi: S \to \mathscr{Y}$ 是映射.

(1) 若对任意的 $\lambda \in (0, 1)$ 有

$$\lambda\varphi(x^1) + (1-\lambda)\varphi(x^2) - \varphi(\lambda x^1 + (1-\lambda)x^2) \in K \quad \forall x^1, x^2 \in S, \tag{2.4.1}$$

则称 φ 是集合 S 上的 K-凸映射,或映射 φ 在 S 上是 K-凸的.

(2) 若对任意的 $\lambda \in (0, 1)$ 有

$$\lambda\varphi(x^1) + (1-\lambda)\varphi(x^2) - \varphi(\lambda x^1 + (1-\lambda)x^2) \in \mathrm{int}K\big(K\backslash\{\boldsymbol{0}\}\big) \quad \forall x^1, x^2 \in S,\ \varphi(x^1) \neq \varphi(x^2), \tag{2.4.2}$$

则称 φ 是集合 S 上的 K-严格凸映射(K-次严格凸映射),或映射 φ 在 S 上是 K-严格凸的(K-次严格凸的).

(3) 若对任意的 $\lambda \in (0, 1)$ 有

$$\lambda\varphi(x^1) + (1-\lambda)\varphi(x^2) - \varphi\big(\lambda x^1 + (1-\lambda)x^2\big) \in \mathrm{int}K \quad \forall x^1, x^2 \in S,\ x^1 \neq x^2, \tag{2.4.3}$$

则称 φ 是集合 S 上的 K-强凸映射,或映射 φ 在 S 上是 K-强凸的.

注 2.4.1 由定义 2.4.1 直接可知，凸集上的 K-强凸映射也是 K-凸的，K-严格凸的和 K-次严格凸的，反之不然. 但是，K-严格凸映射却不一定是 K-凸映射.

注 2.4.2 特别地，在定义 2.4.1 中，当 $\mathscr{Y}=R^m$ 和 $K=R_+^m$ $(m\geqslant 1)$ 时，$\varphi=\boldsymbol{f}:S\to R^m$ 是向量函数. $\boldsymbol{f}$ 在 S 上是 R_+^m-凸的，即对任意的 $\lambda\in(0,1)$ 有

$$\lambda \boldsymbol{f}(\boldsymbol{x}^1)+(1-\lambda)\boldsymbol{f}(\boldsymbol{x}^2)-\boldsymbol{f}(\lambda \boldsymbol{x}^1+(1-\lambda)\boldsymbol{x}^2)\in R_+^m,$$

或即

$$\boldsymbol{f}(\lambda \boldsymbol{x}^1+(1-\lambda)\boldsymbol{x}^2)$$

$$\leqslant \lambda \boldsymbol{f}(\boldsymbol{x}^1)+(1-\lambda)\boldsymbol{f}(\boldsymbol{x}^2)\quad \forall \boldsymbol{x}^1,\boldsymbol{x}^2\in S. \qquad (2.4.4)$$

这时，称 $\boldsymbol{f}$ 为 S 上的凸向量函数. 将上述向量不等式(2.4.4)与(2.1.1)对比可知，$\boldsymbol{f}$ 是 S 上的凸向量函数也即它的每一分量函数在 S 上是凸的. 同样地，$\boldsymbol{f}:S\to R^m$ 在 S 上是 R_+^m-严格凸的(R_+^m-强凸的)，即对任意的 $\lambda\in(0,1)$ 有

$$\boldsymbol{f}(\lambda \boldsymbol{x}^1+(1-\lambda)\boldsymbol{x}^2)<\lambda \boldsymbol{f}(\boldsymbol{x}^1)+(1-\lambda)\boldsymbol{f}(\boldsymbol{x}^2)$$

$$\forall \boldsymbol{x}^1,\boldsymbol{x}^2\in S,\ \boldsymbol{f}(\boldsymbol{x}^1)\neq \boldsymbol{f}(\boldsymbol{x}^2)\ (\boldsymbol{x}^1\neq \boldsymbol{x}^2),$$

并称 $\boldsymbol{f}$ 是 S 上的严格(强)凸向量函数. 这时，$\boldsymbol{f}$ 的每一分量函数在 S 上是严格(强)凸的. 显然，$\boldsymbol{f}:S\to R^m$ 在 S 上是 R_+^m-次严格凸的，也即它的每一分量函数是凸的，并且至少有一是严格凸的. 若 $\mathscr{Y}=R$ 和 $K=R_+$，则由定义 2.4.1 和定义 2.1.1 可知，$f:S\to R$ 是 R_+-凸的，即 f 是 S 上的凸函数，$f:S\to R$ 是 R_+-严格凸的和 R_+-次严格凸的，即 f 是 S 上的严格凸函数，而 $f:S\to R$ 是 R_+-强凸的，即 f 是 S 上的强凸函数.

以下给出有关 K-凸映射的几个基本关系.

定理 2.4.1 设 $S\subset\mathscr{X}$ 是非空凸集，$\varphi:S\to\mathscr{Y}$ 是映射.

(1) 若 φ 是 S 上的 K-凸映射，则 $\varphi(S)\subset\mathscr{Y}$ 是 K-凸集.

(2) 若 φ 是 S 上的 K-强凸映射，则 $\varphi(S)\subset\mathscr{Y}$ 是严格 K-

凸集.

证明　(1) 由 $S\neq\varnothing$ 知 $\varphi(S)\neq\varnothing$. 若 $\varphi(S)$ 是单点集，由定义 1.4.4 可知它是 K-凸集. 现设 $\varphi(S)$ 至少包含两个点，任取 $\boldsymbol{y}^1$，$\boldsymbol{y}^2\in\varphi(S)$，则存在 $\boldsymbol{x}^1$，$\boldsymbol{x}^2\in S$，使

$$\boldsymbol{y}^1=\varphi(\boldsymbol{x}^1),\ \boldsymbol{y}^2=\varphi(\boldsymbol{x}^2). \tag{2.4.5}$$

因为 S 是凸集，故对任意的 $\lambda\in(0,1)$ 有

$$\lambda\boldsymbol{x}^1+(1-\lambda)\boldsymbol{x}^2\in S. \tag{2.4.6}$$

又 φ 是 S 上的 K-凸映射，由定义 2.4.1 的(1)知

$$\lambda\varphi(\boldsymbol{x}^1)+(1-\lambda)\varphi(\boldsymbol{x}^2)-\varphi\left(\lambda\boldsymbol{x}^1+(1-\lambda)\boldsymbol{x}^2\right)\in K.$$

据此，由(2.4.5)和(2.4.6)得到

$$\lambda\boldsymbol{y}^1+(1-\lambda)\boldsymbol{y}^2\in\varphi(S)+K,$$

所以按定义 1.4.4 便知 $\varphi(S)$ 是 K-凸的.

(2) 与(1)的证明类似. 利用定义 2.4.1 的(3)和定义 1.4.5 即可得证. ▯

定理 2.4.2　设 $S\subset\mathscr{X}$ 是非空凸集，则 $\varphi: S\to\mathscr{Y}$ 是 K-凸映射当且仅当对任何正整数 $m\geqslant 2$ 和任意的 $\lambda_i\geqslant 0$ $(i=1,\cdots,m)$，$\sum\limits_{i=1}^{m}\lambda_i=1$，有

$$\sum_{i=1}^{m}\lambda_i\varphi(\boldsymbol{x}^i)-\varphi\left(\sum_{i=1}^{m}\lambda_i\boldsymbol{x}^i\right)\in K\quad\forall\boldsymbol{x}^1,\cdots,\boldsymbol{x}^m\in S. \tag{2.4.7}$$

证明　充分性. 在(2.4.7)中取 $m=2$，由定义 2.4.1 中的(1)即知 φ 是 S 上的 K-凸映射.

必要性. 用数学归纳法，与定理 2.1.1 证明中的必要性部分类似，由定义 2.4.1 的(1)注意到 K 是凸锥即可得证. ▯

定理 2.4.3　设 $S\subset\mathscr{X}$ 是非空凸集. 若 $\varphi: S\to\mathscr{Y}$ 是 K-凸映

射，则对任意的 $c \in \mathscr{Y}$，K-水平集 $H_S(\varphi, c)_K = \{x \in S \mid c - \varphi(x) \in K\}$ 是凸集.

证明　任取 $x^1, x^2 \in H_S(\varphi, c)_K \subset S$，则有

$$c - \varphi(x^1),\ c - \varphi(x^2) \in K. \tag{2.4.8}$$

因为 S 是凸集，故对任意的 $\lambda \in (0, 1)$，有 $\lambda x^1 + (1-\lambda)x^2 \in S$. 由于 φ 是 S 上的 K-凸映射，按定义 2.4.1 的(1)知

$$\lambda\varphi(x^1) + (1-\lambda)\varphi(x^2) - \varphi(\lambda x^1 + (1-\lambda)x^2) \in K.$$

据此，由(2.4.8)注意到 K 是凸锥，有

$$\begin{aligned} c - \varphi(\lambda x^1 + (1-\lambda)x^2) &\in c + K - \lambda\varphi(x^1) - (1-\lambda)\varphi(x^2) \\ &= K + \lambda[c - \varphi(x^1)] + (1-\lambda)[c - \varphi(x^2)] \\ &\subset K + \lambda K + (1-\lambda)K \\ &= K + K \subset K. \end{aligned}$$

于是，得到 $\lambda x^1 + (1-\lambda)x^2 \in H_S(\varphi, c)_K$，因而由定义 1.1.1 即知 $H_S(\varphi, c)_K$ 是凸集. ▯

注 2.4.3　当 $\mathscr{Y} = R$ 和 $K = R_+$ 时，$\varphi = f: S \to R$ 是实值函数，R_+-水平集即水平集：

$$\begin{aligned} H_S(f, c)_{R_+} &= \{x \in S \mid c - f(x) \in R_+\} \\ &= \{x \in S \mid f(x) \leqslant c\} = H_S(f, c). \end{aligned}$$

这时，定理 2.4.2 和定理 2.4.3 分别即是定理 2.1.1 和定理 2.1.2.

为讨论 K-凸映射和 K-严格凸映射之间的关系，下面引入映射的 K-半连续性概念. 设 $\mathscr{X}$ 是线性拓扑空间.

定义 2.4.2　设集合 $S \subset \mathscr{X}$ 非空，$\varphi: S \to \mathscr{Y}$ 是映射，$K \subset \mathscr{Y}$ 是非平凡锥. 若对任意的 $c \in \mathscr{Y}$，K-水平集

$$H_S(\varphi, c)_K = \{x \in S \mid c - \varphi(x) \in \mathrm{cl}K\}$$

是闭集,则称 φ 是 S 上的 K-半连续映射,或映射 φ 在 S 上是 K-半连续的.

注 2.4.4 设 $\mathscr{Y}=R$ 和 $K=R_+$,由定义 2.4.2 可知 $\varphi=f$ 在 S 上是 R_+-半连续的,意味着 R_+-水平集 $H_S(f, c)_{R_+}=\{x\in S\,|\,c-f(x)\in R_+\}$,也即水平集 $H_S(f, c)=\{x\in S\,|\,f(x)\leqslant c\}$ 是闭的.由定理 2.2.5 的(1)和(3),这就是函数 f 在 S 上是下半连续的.同理,f 在 S 上是 R_--半连续的,意即 $\{x\in S\,|\,f(x)\geqslant c\}$ 是闭集,于是由定理 2.2.6 的(1)和(3),也即函数 f 在 S 上是上半连续的.

定理 2.4.4 设 $S\subset\mathscr{X}$ 是非空凸集.若 $\varphi: S\to\mathscr{Y}$ 是 K-半连续的 K-严格凸映射,则 φ 是 S 上的 K-凸映射.

证明 设对任意的 $\boldsymbol{x}^1, \boldsymbol{x}^2\in S$ 有 $\varphi(\boldsymbol{x}^1)\neq\varphi(\boldsymbol{x}^2)$.由 φ 在 S 上是 K-严格凸映射,按定义 2.4.1 的(2)有

$$\lambda\varphi(\boldsymbol{x}^1)+(1-\lambda)\varphi(\boldsymbol{x}^2)-\varphi(\lambda\boldsymbol{x}^1+(1-\lambda)\boldsymbol{x}^2)\in\operatorname{int}K\subset K.$$

因此,由定义 2.4.1 的(1),即知 φ 是 K-凸映射.

设存在 $\bar{\boldsymbol{x}}^1, \bar{\boldsymbol{x}}^2\in S$ 使 $\varphi(\bar{\boldsymbol{x}}^1)=\varphi(\bar{\boldsymbol{x}}^2)$.用反证法,假设 φ 在 S 上不是 K-凸映射,则按定义 2.4.1 的(1)可知,存在 $\lambda_0\in(0, 1)$,有

$$\lambda_0\varphi(\bar{\boldsymbol{x}}^1)+(1-\lambda_0)\varphi(\bar{\boldsymbol{x}}^2)-\varphi(\lambda_0\bar{\boldsymbol{x}}^1+(1-\lambda_0)\bar{\boldsymbol{x}}^2)\notin K.$$

令 $\boldsymbol{x}^0=\lambda_0\bar{\boldsymbol{x}}^1+(1-\lambda_0)\bar{\boldsymbol{x}}^2$,则得

$$\varphi(\bar{\boldsymbol{x}}^1)-\varphi(\boldsymbol{x}^0)\notin K. \tag{2.4.9}$$

记

$$\lambda'=\inf\{\lambda\in(\lambda_0, 1]\,|\,\varphi(\bar{\boldsymbol{x}}^1)-\varphi(\lambda\bar{\boldsymbol{x}}^1+(1-\lambda)\bar{\boldsymbol{x}}^2)\in K\}, \tag{2.4.10}$$

则有 $\lambda'>\lambda_0$(否则,若有 $\lambda'=\lambda_0$,由 φ 在 S 上是 K-半连续的,按定义 2.4.2 知 $H_S(\varphi, \varphi(\bar{\boldsymbol{x}}^1))_K=\{\boldsymbol{x}\in S\,|\,\varphi(\bar{\boldsymbol{x}}^1)-\varphi(\boldsymbol{x})\in K\}$ 是闭集.由此,设 $\{\lambda_k\}\subset(\lambda_0, 1]$, $\lambda_k\bar{\boldsymbol{x}}^1+(1-\lambda_k)\bar{\boldsymbol{x}}^2\in H_S(\varphi, \varphi(\bar{\boldsymbol{x}}^1))_K$, $\lambda_k\to\lambda'=\lambda_0(k\to\infty)$,则有 $\lambda_k\bar{\boldsymbol{x}}^1+(1-\lambda_k)\bar{\boldsymbol{x}}^2\to\lambda_0\bar{\boldsymbol{x}}^1+(1-\lambda_0)\bar{\boldsymbol{x}}^2$

$\in H_S(\varphi, \varphi(\bar{x}^1))_K$，故得 $\varphi(\bar{x}^1) - \varphi(x^0) \in K$，这与(2.4.9)相矛盾). 现令

$$x' = \lambda' \bar{x}^1 + (1 - \lambda') \bar{x}^2,$$

由(2.4.10)和 $\lambda' > \lambda_0$ 知

$$\varphi(\bar{x}^1) - \varphi(x) \notin K \quad \forall x \in (x', x^0). \qquad (2.4.11)$$

任取 $\tilde{x} \in (x', x^0)$，则 $\tilde{x} \in (\bar{x}^1, x^0)$. 记 $\tilde{x} = \mu \bar{x}^1 + (1 - \mu) x^0 (\mu \in (0, 1))$，由(2.4.9)有 $\varphi(\bar{x}^1) \neq \varphi(x^0)$，因为 φ 是 K-严格凸映射，据定义 2.4.1 的(2)知

$$\varphi(\tilde{x}) \in \mu\varphi(\bar{x}^1) + (1 - \mu)\varphi(x^0) - \text{int}K. \qquad (2.4.12)$$

另外，因为 $x^0 \in (\tilde{x}, \bar{x}^2)$，记 $x^0 = \nu \tilde{x} + (1 - \nu) \bar{x}^2 (\nu \in (0, 1))$，从(2.4.11)有 $\varphi(\tilde{x}) \neq \varphi(\bar{x}^1) = \varphi(\bar{x}^2)$，所以由 φ 是 K-严格凸的又知

$$\varphi(x^0) \in \nu\varphi(\tilde{x}) + (1 - \nu)\varphi(\bar{x}^2) - \text{int}K,$$

即

$$\varphi(x^0) \in \nu\varphi(\tilde{x}) + (1 - \nu)\varphi(\bar{x}^1) - \text{int}K. \qquad (2.4.13)$$

作 ν(2.4.12) + (2.4.13)，有

$$\nu\varphi(\tilde{x}) + \varphi(x^0) \in \left[\nu\mu\varphi(\bar{x}^1) + \nu(1 - \mu)\varphi(x^0) - \text{int}K\right] + \left[\nu\varphi(\tilde{x}) + (1 - \nu)\varphi(\bar{x}^1) - \text{int}K\right],$$

即

$$(1 - \nu + \nu\mu)\varphi(x^0) \in (1 - \nu + \nu\mu)\varphi(\bar{x}^1) - \text{int}K,$$

从而得 $\varphi(\bar{x}^1) - \varphi(x^0) \in \text{int}K$，这导致与(2.4.9)相矛盾. ▯

注 2.4.5 当 $\mathscr{Y} = R$ 和 $K = R_+$ 时，定理 2.4.4 即定理 2.1.3.

以下阐述锥拟凸映射类.

先引进向量组的锥界集概念，借助它给出锥拟凸映射类的定义.

定义 2.4.3　设 $y^1, \cdots, y^m \in \mathscr{Y}(m \geqslant 2)$. 若 $\bar{y} \in \bigcap_{i=1}^{m}(y^i + K)$, 并且不存在 $y \in \bigcap_{i=1}^{m}(y^i + K)$ 使得

$$\bar{y} - y \in K \backslash \{\boldsymbol{0}\},$$

则称 $\bar{y}$ 是向量组 $\{y^1, \cdots, y^m\}$ 的 K-界点. $\{y^1, \cdots, y^m\}$ 的所有 K-界点组成的集合称为它的 K-界集, 记作 $K\text{-bou}\{y^1, \cdots, y^m\}$.

注 2.4.6　对于给定的 $y^1, \cdots, y^m \in \mathscr{Y}$, 由 $\text{int}K \neq \varnothing$ 可以推得 $\bigcap_{i=1}^{m}(y^i + K) \neq \varnothing$, 并且由 K 是尖闭凸锥不难得知 $\bigcap_{i=1}^{m}(y^i + K)$ 是 K-有界和 K-闭的(见文献[27]的定义 3.3.2). 据此, 由文献[27]的定理 3.3.6 可知 $\{y^1, \cdots, y^m\}$ 的 K-界集是存在的.

定义 2.4.4　设 $S \subset \mathscr{X}$ 是非空凸集, $\varphi: S \to \mathscr{Y}$ 是映射.

(1) 若对任意的 $\lambda \in (0, 1)$ 有

$$K - \text{bou}\{\varphi(x^1), \varphi(x^2)\} - \varphi(\lambda x^1 + (1-\lambda)x^2) \subset K \quad \forall x^1, x^2 \in S, \tag{2.4.14}$$

则称 φ 是集合 S 上的 K-拟凸映射, 或映射 φ 在 S 上是 K-拟凸的.

(2) 若对任意的 $\lambda \in (0, 1)$ 有

$$K - \text{bou}\{\varphi(x^1), \varphi(x^2)\} - \varphi(\lambda x^1 + (1-\lambda)x^2) \subset \text{int}K\,(K\backslash\{\boldsymbol{0}\}) \quad \forall x^1, x^2 \in S, \varphi(x^1) \neq \varphi(x^2), \tag{2.4.15}$$

则称 φ 是集合 S 上的 K-严格拟凸映射(K-次严格拟凸映射), 或映射 φ 在 S 上是 K-严格拟凸的(K-次严格拟凸的).

(3) 若对任意的 $\lambda \in (0, 1)$ 有

$$K - \text{bou}\{\varphi(x^1), \varphi(x^2)\} - \varphi(\lambda x^1 + (1-\lambda)x^2) \subset \text{int}K \quad \forall x^1, x^2 \in S, x^1 \neq x^2, \tag{2.4.16}$$

则称 φ 是集合 S 上的 K-强拟凸映射, 或映射 φ 在 S 上是 K-强拟凸的.

注 2.4.7　从定义 2.4.4 直接可以得知，凸集上的 K-强拟凸映射也是 K-拟凸的、K-严格拟凸的和 K-次严格拟凸的，又 K-严格拟凸映射也是 K-次严格拟凸的，反之不然. 然而，K-严格拟凸映射却不一定是 K-拟凸映射.

注 2.4.8　在定义 2.4.4 中当 $\mathscr{Y} = R$ 和 $K = R_+$ 时，$\varphi = f: S \to \mathscr{Y}$ 是实值函数. 这时，因为

$$\bigcap_{i=1}^{2}\left(f(\boldsymbol{x}^i) + R_+\right) = \left[\max\{f(\boldsymbol{x}^1),\ f(\boldsymbol{x}^2)\},\ +\infty\right),$$

由定义 2.4.3 可知

$$R_+ - \operatorname{bou}\{f(\boldsymbol{x}^1),\ f(\boldsymbol{x}^2)\} = \max\{f(\boldsymbol{x}^1),\ f(\boldsymbol{x}^2)\}.$$

于是，定义 2.4.4 中的(2.4.14)即

$$f(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \leqslant \max\{f(\boldsymbol{x}^1),\ f(\boldsymbol{x}^2)\}$$

$$\forall \boldsymbol{x}^1,\ \boldsymbol{x}^2 \in S.$$

对照定义 2.3.1 中的(2.3.1)，便知 R_+-拟凸映射即拟凸函数. 同理，R_+-严格拟凸映射和 R_+-次严格拟凸映射(这时二者等同)即严格拟凸函数，R_+-强拟凸映射即强拟凸函数.

注 2.4.9　在定义 2.4.4 中当 $\mathscr{Y} = R^m$ 和 $K = R_+^m\ (m \geqslant 2)$ 时，$\varphi = f: S \to R^m$ 是向量函数. 设 $\boldsymbol{f}(\boldsymbol{x}^i) = (f_1(\boldsymbol{x}^i),\ \cdots,\ f_m(\boldsymbol{x}^i))^T$ $(i = 1,\ 2)$，则

$$\bigcap_{i=1}^{2}(f(\boldsymbol{x}^i) + R_+^m) = \begin{bmatrix} \max\{f_1(\boldsymbol{x}^1),\ f_1(\boldsymbol{x}^2)\} \\ \vdots \\ \max\{f_m(\boldsymbol{x}^1),\ f_m(\boldsymbol{x}^2)\} \end{bmatrix} + R_+^m,$$

于是由定义 2.4.3 可得知

$$R_+^m - \operatorname{bou}\{f(\boldsymbol{x}^1),\ f(\boldsymbol{x}^2)\} = \begin{bmatrix} \max\{f_1(\boldsymbol{x}^1),\ f_1(\boldsymbol{x}^2)\} \\ \vdots \\ \max\{f_m(\boldsymbol{x}^1),\ f_m(\boldsymbol{x}^2)\} \end{bmatrix}.$$

据此，由定义 2.4.4 的(1)可知，$\boldsymbol{f}$ 是 S 上的 R^m_+-拟凸映射即对任意的 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in S$ 和任意的 $\lambda \in (0, 1)$ 有

$$f_i(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \leqslant \max\{f_i(\boldsymbol{x}^1), f_i(\boldsymbol{x}^2)\}, i = 1, \cdots, m.$$

对照(2.3.1)得知，这时 $\boldsymbol{f}$ 的每一分量函数均是 S 上的拟凸函数。同理，$\boldsymbol{f}$ 是 R^m_+-严格拟凸映射和 R^m_+-强拟凸映射也即它们的分量函数分别都是严格拟凸函数和强拟凸函数，而 $\boldsymbol{f}$ 是 R^m_+-次严格拟凸映射即它的每一分量都是拟凸函数并且其中至少有一个是严格拟凸函数.

锥凸映射类和锥拟凸映射类有如下相应关系.

定理 2.4.5 设 $S \subset \mathscr{X}$ 是非空凸集，φ: $S \to \mathscr{Y}$ 是映射.

(1) 若 φ 是 S 上的 K-凸映射，则 φ 是 S 上的 K-拟凸映射.

(2) 若 φ 是 S 上的 K-严格凸映射(K-次严格凸映射)，则 φ 是 S 上的 K-严格拟凸映射(K-次严格拟凸映射).

(3) 若 φ 是 S 上的 K-强凸映射，则 φ 是 S 上的 K-强拟凸映射.

证明 (1) 由于 φ 在 S 上是 K-凸的，按定义 2.4.1 的(1)可知，对任意的 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in S$ 和任意的 $\lambda \in (0, 1)$ 有

$$\lambda\varphi(\boldsymbol{x}^1) + (1-\lambda)\varphi(\boldsymbol{x}^2) - \varphi(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \in K,$$

或即

$$\varphi(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \in \lambda\varphi(\boldsymbol{x}^1) + (1-\lambda)\varphi(\boldsymbol{x}^2) - K. \tag{2.4.17}$$

依据定义 2.4.3，对任何 $\bar{\boldsymbol{y}} \in K - \mathrm{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\}$ 有 $\bar{\boldsymbol{y}} \in \bigcap_{i=1}^{2}(\varphi(\boldsymbol{x}^i) + K)$，因而

$$K - \mathrm{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\} \subset \varphi(\boldsymbol{x}^i) + K, i = 1, 2,$$

故

$$\varphi(\boldsymbol{x}^i) \in K - \mathrm{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\} - K, i = 1, 2.$$

据此，由(2.4.17)得

$$\varphi\left(\lambda x^1+(1-\lambda)x^2\right)\in\lambda\left(K-\mathrm{bou}\{\varphi(x^1),\varphi(x^2)\}-K\right)$$
$$+(1-\lambda)\left(K-\mathrm{bou}\{\varphi(x^1),\right.$$
$$\left.\varphi(x^2)\}-K\right)-K$$
$$=K-\mathrm{bou}\{\varphi(x^1),\varphi(x^2)\}-K-K$$
$$\subset K-\mathrm{bou}\{\varphi(x^1),\varphi(x^2)\}-K,$$

从而

$$K-\mathrm{bou}\{\varphi(x^1),\varphi(x^2)\}-\varphi\left(\lambda x^1+(1-\lambda)x^2\right)\subset K.$$

因此，由定义 2.4.4 的(1)，即知 φ 在 S 上是 K-拟凸映射.

(2)和(3)与(1)的证明类似. □

以下给出有关锥拟凸映射的两个基本结果.

定理 2.4.6 设 $S\subset\mathscr{X}$ 是非空凸集，则 $\varphi: S\to\mathscr{Y}$ 是 K-拟凸映射当且仅当对任何正整数 $m\geqslant 2$ 和任意的 $\lambda_i\geqslant 0\ (i=1,\cdots,m)$, $\sum\limits_{i=1}^{m}\lambda_i=1$，有

$$K-\mathrm{bou}\{\varphi(x^1),\cdots,\varphi(x^m)\}-\varphi\left(\sum_{i=1}^{m}\lambda_i x^i\right)\subset K$$
$$\forall x^1,\cdots,x^m\in S. \tag{2.4.18}$$

证明 充分性. 在(2.4.18)中取 $m=2$，由定义 2.4.4 中的(1)可知 φ 是 S 上的 K-拟凸映射.

必要性. 用数学归纳法证明，当 $m=2$ 时，从(2.4.14)知(2.4.18)成立. 现设 $m\leqslant k$ ($k\geqslant 2$ 是正整数)时(2.4.18)成立，以下证明(2.4.18)对于 $m=k+1$ 成立. 不失一般性，设 $\lambda_i>0$ ($i=1,\cdots,k+1$), $\sum\limits_{i=1}^{k+1}\lambda_i=1$ (若有某个 $\lambda_i=0$，则 $m\leqslant k$，由假设(2.4.18)成立)，则有 $1-\lambda_{k+1}=\sum\limits_{i=1}^{k}\lambda_i>0$. 记 $\sum\limits_{i=1}^{k}\lambda_i x^i=(1-$

$\lambda_{k+1})\boldsymbol{x}$，由 $\boldsymbol{x}^i \in S$ 和 S 是凸集，有 $\boldsymbol{x} = \sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}}\boldsymbol{x}^i \in S$. 由于已知 φ 在 S 上是 K-拟凸的，按定义 2.4.4 的(1)得知

$$K - \mathrm{bou}\{\varphi(\boldsymbol{x}), \varphi(\boldsymbol{x}^{k+1})\} - \varphi(\lambda_{k+1}\boldsymbol{x}^{k+1} + (1-\lambda_{k+1})\boldsymbol{x}) \subset K,$$

或即

$$K - \mathrm{bou}\left\{\varphi\left(\sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}}\boldsymbol{x}^i\right), \varphi(\boldsymbol{x}^{k+1})\right\} \subset \varphi\left(\sum_{i=1}^{k+1}\lambda_i \boldsymbol{x}^i\right) + K. \tag{2.4.19}$$

另外，由 $\sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}} = 1$ 和归纳法假设(2.4.18)对于 $m = k$ 成立，因而

$$K - \mathrm{bou}\{\varphi(\boldsymbol{x}^1), \cdots, \varphi(\boldsymbol{x}^k)\} - \varphi\left(\sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}}\boldsymbol{x}^i\right) \subset K. \tag{2.4.20}$$

因为显然有 $\bigcap_{i=1}^{k+1}(\varphi(\boldsymbol{x}^i) + K) \subset \bigcap_{i=1}^{k}(\varphi(\boldsymbol{x}^i) + K)$，所以由定义 2.4.3 可以推知

$$\begin{aligned} &K - \mathrm{bou}\{\varphi(\boldsymbol{x}^1), \cdots, \varphi(\boldsymbol{x}^k), \varphi(\boldsymbol{x}^{k+1})\} \\ &\subset K - \mathrm{bou}\{\varphi(\boldsymbol{x}^1), \cdots, \varphi(\boldsymbol{x}^k)\} + K. \end{aligned}$$

利用(2.4.20)并注意 K 是凸锥，得到

$$K - \mathrm{bou}\{\varphi(\boldsymbol{x}^1), \cdots, \varphi(\boldsymbol{x}^{k+1})\} \subset \varphi\left(\sum_{i=1}^{k} \frac{\lambda_i}{1-\lambda_{k+1}}\boldsymbol{x}^i\right) + K.$$

由定义 2.4.3 知

$$\begin{aligned} K - \mathrm{bou}\{\varphi(\boldsymbol{x}^1), \cdots, \varphi(\boldsymbol{x}^{k+1})\} &\subset \bigcap_{i=1}^{k+1}(\varphi(\boldsymbol{x}^i) + K) \\ &\subset \varphi(\boldsymbol{x}^{k+1}) + K, \end{aligned}$$

从而

$$K-\text{bou}\{\varphi(\boldsymbol{x}^1),\cdots,\varphi(\boldsymbol{x}^{k+1})\}$$

$$\subset\left(\varphi\left(\sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}}\boldsymbol{x}^i\right)+K\right)\cap\left(\varphi(\boldsymbol{x}^{k+1})+K\right).$$

又从定义 2.4.3 易知

$$\left(\varphi\left(\sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}}\boldsymbol{x}^i\right)+K\right)\cap(\varphi(\boldsymbol{x}^{k+1})+K)$$

$$\subset K-\text{bou}\left\{\varphi\left(\sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}}\boldsymbol{x}^i\right),\varphi(\boldsymbol{x}^{k+1})\right\}+K,$$

因此有

$$K-\text{bou}\{\varphi(\boldsymbol{x}^1),\cdots,\varphi(\boldsymbol{x}^{k+1})\}$$

$$\subset K-\text{bou}\left\{\varphi\left(\sum_{i=1}^{k}\frac{\lambda_i}{1-\lambda_{k+1}}\boldsymbol{x}^i\right),\varphi(\boldsymbol{x}^{k+1})\right\}+K.$$

最后,由(2.4.19)并注意到 K 是凸锥,即得

$$K-\text{bou}\{\varphi(\boldsymbol{x}^1),\cdots,\varphi(\boldsymbol{x}^{k+1})\}\subset\varphi\left(\sum_{i=1}^{k+1}\lambda_i\boldsymbol{x}^i\right)+K,$$

从而(2.4.18)对于 $m=k+1$ 成立. □

定理 2.4.7 设 $S\subset\mathscr{X}$ 是非空凸集,则 $\varphi: S\to\mathscr{Y}$ 是 K-拟凸映射当且仅当对任意的 $\boldsymbol{c}\in\mathscr{Y}$, K-水平集 $H_S(\varphi,\boldsymbol{c})_K=\{\boldsymbol{x}\in S\mid\boldsymbol{c}-\varphi(\boldsymbol{x})\in K\}$ 是凸集.

证明 必要性. 对任意的 $\boldsymbol{c}\in\mathscr{Y}$, 任取 $\boldsymbol{x}^1,\boldsymbol{x}^2\in H_S(\varphi,\boldsymbol{c})_K\subset S$, 则有 $\boldsymbol{c}-\varphi(\boldsymbol{x}^1)\in K$ 和 $\boldsymbol{c}-\varphi(\boldsymbol{x}^2)\in K$, 从而

$$\boldsymbol{c}\in(\varphi(\boldsymbol{x}^1)+K)\cap(\varphi(\boldsymbol{x}^2)+K).$$

据此,由定义 2.4.3 得知,存在 $\bar{\boldsymbol{y}}\in K-\text{bou}\{\varphi(\boldsymbol{x}^1),\varphi(\boldsymbol{x}^2)\}$ 使 $\boldsymbol{c}\in\bar{\boldsymbol{y}}+K$, 或

$$\boldsymbol{c}-\bar{\boldsymbol{y}}\in K. \tag{2.4.21}$$

由于 φ 在 S 上是 K-拟凸的，由定义 2.4.4 的(1)知，对任意的 $\lambda \in (0, 1)$ 有

$$K - \operatorname{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\} - \varphi(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \subset K,$$

从而有

$$\bar{\boldsymbol{y}} - \varphi(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \in K.$$

于是，注意到(2.4.21)和 K 是凸锥而得到

$$\begin{aligned} &\boldsymbol{c} - \varphi(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \\ &= (\boldsymbol{c} - \bar{\boldsymbol{y}}) + (\bar{\boldsymbol{y}} - \varphi(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2)) \in K + K \subset K, \end{aligned}$$

因而 $\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 \in H_S(\varphi, \boldsymbol{c})_K$. 由此，按定义 1.1.1 知 $H_S(\varphi, \boldsymbol{c})_K$ 是凸集.

充分性. 对任意的 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in S$, 任取 $\boldsymbol{c} \in K - \operatorname{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\}$. 由于显然有 $K - \operatorname{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\} \subset (\varphi(\boldsymbol{x}^1) + K) \cap (\varphi(\boldsymbol{x}^2) + K)$, 故有

$$\boldsymbol{c} - \varphi(\boldsymbol{x}^1) \in K, \ \boldsymbol{c} - \varphi(\boldsymbol{x}^2) \in K,$$

于是 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in H_S(\varphi, \boldsymbol{c})_K$. 因为 $H_S(\varphi, \boldsymbol{c})_K$ 是凸集，故对任意的 $\lambda \in (0, 1)$ 有 $\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 \in H_S(\varphi, \boldsymbol{c})_K$, 从而

$$\boldsymbol{c} - \varphi(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \in K.$$

由 $\boldsymbol{c} \in K - \operatorname{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\}$ 的任意性，得到

$$K - \operatorname{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\} - \varphi(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \subset K,$$

因此由定义 2.4.4 的(1)知 φ 在 S 上是 K-拟凸的. □

定理 2.1.3 和定理 2.3.3 曾经分别给出了在下半连续的条件下，严格凸函数是凸函数和严格拟凸函数是拟凸函数的结果. 以下讨论 K-严格拟凸映射和 K-拟凸映射的类似关系.

设 $\boldsymbol{a} \in \mathscr{Y}$, $\boldsymbol{\omega} \in \operatorname{int} K$, 考虑 Minkowski 泛函

$$\mu_a : \mathscr{Y} \times \operatorname{int} K \to R,\ (y, \omega) \mapsto \mu_a(y, \omega),$$

其中

$$\mu_a(y, \omega) = \inf\{\alpha \in R \mid y \in a + \alpha\omega - K\},$$

$$y \in \mathscr{Y},\ \omega \in \operatorname{int} K. \tag{2.4.22}$$

显然,对于任一 $y \in \mathscr{Y}$, $\mu_a(y, \omega)$ 是一确定的值,并且有

$$y \in a + \mu_a(y, \omega)\omega - K. \tag{2.4.23}$$

引理 2.4.8 设 $y^1, y^2 \in \mathscr{Y}$.

(1) 若 $y^1 - y^2 \in K$,则 $\mu_a(y^1, \omega) \geqslant \mu_a(y^2, \omega)$.

(2) 若 $y^1 - y^2 \in \operatorname{int} K$, 则 $\mu_a(y^1, \omega) > \mu_a(y^2, \omega)$.

证明 (1) 由 $y^1 - y^2 \in K$,则 $y^2 \in y^1 - K$. 对 y^1 利用(2.4.23)有

$$y^2 \in \{a + \mu_a(y^1, \omega)\omega - K\} - K$$
$$\subset a + \mu_a(y^1, \omega)\omega - K.$$

据此,由(2.4.22)得到 $\mu_a(y^2, \omega) \leqslant \mu_a(y^1, \omega)$.

(2) 由 $y^1 - y^2 \in \operatorname{int} K$,对 $\omega \in \operatorname{int} K$ 存在 $\varepsilon > 0$,使 $y^1 - y^2 - \varepsilon\omega \in K$,即 $y^2 \in y^1 - \varepsilon\omega - K$. 利用(2.4.23),有

$$y^2 \in \{a + \mu_a(y^1, \omega)\omega - K\} - \varepsilon\omega - K$$
$$\subset a + \{\mu_a(y^1, \omega) - \varepsilon\}\omega - K.$$

由此,根据(2.4.22)得到 $\mu_a(y^2, \omega) \leqslant \mu_a(y^1, \omega) - \varepsilon < \mu_a(y^1, \omega)$. ▯

引理 2.4.9 设 $a \in \mathscr{Y}$, 则对任意的 $y^1, y^2 \in \mathscr{Y}$ 有

$$\inf\{\mu_a(y, \omega) \mid y \in K - \operatorname{bou}\{y^1, y^2\}\}$$
$$= \max\{\mu_a(y^1, \omega), \mu_a(y^2, \omega)\},$$

并且下确界可达.

证明 任取 $y \in K - \operatorname{bou}\{y^1, y^2\}$, 由定义 2.4.3 可知 $y \in$

$(y^1+K)\cap(y^2+K)$，从而

$$y-y^1\in K,\quad y-y^2\in K.$$

据此，由引理 2.4.8 的(1)，有

$$\mu_a(y,\omega)\geqslant\mu_a(y^1,\omega),\quad \mu_a(y,\omega)\geqslant\mu_a(y^2,\omega),$$

于是

$$\mu_a(y,\omega)\geqslant\max\{\mu_a(y^1,\omega),\ \mu_a(y^2,\omega)\},$$

从而得

$$\begin{aligned}&\inf\{\mu_a(y,\omega)\mid y\in K-\mathrm{bou}\{y^1,y^2\}\}\\&\geqslant\max\{\mu_a(y^1,\omega),\ \mu_a(y^2,\omega)\}.\end{aligned}\tag{2.4.24}$$

另一方面，不妨设 $\mu_a(y^1,\omega)\geqslant\mu_a(y^2,\omega)$，则有

$$y^1\in a+\mu_a(y^1,\omega)\omega-K$$

和

$$y^2\in a+\mu_a(y^2,\omega)\omega-K\subset a+\mu_a(y^1,\omega)\omega-K,$$

从而

$$a+\mu_a(y^1,\omega)\omega\in(y^1+K)\cap(y^2+K).$$

据此，由定义 2.4.3 可知存在 $\bar{y}\in K-\mathrm{bou}\{y^1,y^2\}$，使 $a+\mu_a(y^1,\omega)\omega\in\bar{y}+K$，也即

$$\bar{y}\in a+\mu_a(y^1,\omega)\omega-K.$$

根据(2.4.22)知

$$\mu_a(\bar{y},\omega)\leqslant\mu_a(y^1,\omega)=\max\{\mu_a(y^1,\omega),\ \mu_a(y^2,\omega)\},$$

再由(2.4.24)有

$$\begin{aligned}\inf\{\mu_a(y,\omega)\mid y\in K-\mathrm{bou}\{y^1,y^2\}\}&\leqslant\mu_a(\bar{y},\omega)\\&\leqslant\max\{\mu_a(y^1,\omega),\ \mu_a(y^2,\omega)\}\\&\leqslant\inf\{\mu_a(y,\omega)\mid y\in K-\mathrm{bou}\{y^1,y^2\}\},\end{aligned}$$

因而得到

$$\inf\{\mu_a(y, \omega) \mid y \in K - \text{bou}\{y^1, y^2\}\}$$
$$= \max\{\mu_a(y^1, \omega), \mu_a(y^2, \omega)\}.$$

显然，其下确界在 $\bar{y} \in K - \text{bou}\{y^1, y^2\}$ 处达到. ▯

设集合 $S \subset \mathscr{X}$，考虑由映射 $\varphi: S \to \mathscr{Y}$ 确定的实值函数（泛函）$f_a: S \to R$：

$$f_a(x) = \mu_a(\varphi(x), \omega), \ x \in S. \tag{2.4.25}$$

引理 2.4.10 设集合 $S \subset \mathscr{X}$ 非空，则 $\varphi: S \to \mathscr{Y}$ 是 K-半连续的当且仅当对任意的 $a \in \mathscr{Y}$，$f_a: S \to R$ 是下半连续的.

证明 任取 $c \in R$，由(2.4.22)可知对任意的 $a \in \mathscr{Y}$，$\mu_a(y, \omega) \leqslant c$ 意味着 $y \in a + c\omega - K$. 因此，由(2.4.25)有

$$\{x \in S \mid f_a(x) \leqslant c\} = \{x \in S \mid \varphi(x) \in a + c\omega - K\}. \tag{2.4.26}$$

设 φ 在 S 上是 K-半连续的，由定义 2.4.2 知(2.4.26)右端的集合是闭集. 据此，(2.4.26)左端的集合是闭集，由定理 2.2.5 的(1)和(3)即知 f_a 是下半连续的. 反之，若对任意的 $a \in \mathscr{Y}$，f_a 是下半连续的，由定理 2.2.5 的(1)和(3)知 $\{x \in S \mid f_a(x) \leqslant 0\}$ 是闭集. 再根据(2.4.26)，$\{x \in S \mid a - \varphi(x) \in K\}$ 是闭集，于是由定义 2.4.2 得知 φ 是 K-半连续的. ▯

引理 2.4.11 设集合 $S \subset \mathscr{X}$ 非空，$\varphi: S \to \mathscr{Y}$ 是映射.

(1) φ 在 S 上是 K-拟凸的当且仅当对任意的 $a \in \mathscr{Y}$，f_a 在 S 上是拟凸的.

(2) 若 φ 在 S 上是 K-严格拟凸的，则对任意的 $a \in \mathscr{Y}$，f_a 在 S 上是严格拟凸的.

证明 (1) 必要性. 从 φ 在 S 上是 K-拟凸的，由定义 2.4.4 的(1)可知，对任意的 $x^1, x^2 \in S$ 和任意的 $\lambda \in (0, 1)$ 有

$$K - \text{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\} - \varphi(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \subset K. \tag{2.4.27}$$

任取 $\boldsymbol{a} \in \mathscr{Y}$. 由引理 2.4.9 和(2.4.25)知存在 $\bar{\boldsymbol{y}} \in K - \text{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\}$，使

$$\mu_a(\bar{\boldsymbol{y}}, \boldsymbol{\omega}) = \max\{f_a(\boldsymbol{x}^1), f_a(\boldsymbol{x}^2)\}. \tag{2.4.28}$$

根据 $\bar{\boldsymbol{y}} \in K - \text{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\}$ 和(2.4.27)，则 $\bar{\boldsymbol{y}} - \varphi(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \in K$，由引理 2.4.8 的(1)知

$$\mu_a(\bar{\boldsymbol{y}}, \boldsymbol{\omega}) \geqslant \mu_a(\varphi(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2), \boldsymbol{\omega})$$
$$= f_a(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2).$$

于是，由(2.4.28)得到

$$f_a(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \leqslant \max\{f_a(\boldsymbol{x}^1), f_a(\boldsymbol{x}^2)\},$$

从而按定义 2.3.1 的(1)即知 f_a 在 S 上是拟凸的.

充分性. 反之，假设 φ 在 S 上不是 K-拟凸的，根据定义 2.4.4 的(1)，则存在 $\bar{\boldsymbol{x}}^1, \bar{\boldsymbol{x}}^2 \in S$ 和 $\bar{\lambda} \in (0, 1)$ 使

$$K - \text{bou}\{\varphi(\bar{\boldsymbol{x}}^1), \varphi(\bar{\boldsymbol{x}}^2)\} - \varphi(\bar{\lambda}\bar{\boldsymbol{x}}^1 + (1-\bar{\lambda})\bar{\boldsymbol{x}}^2) \not\subset K.$$

于是，有 $\boldsymbol{a}^0 \in K - \text{bou}\{\varphi(\bar{\boldsymbol{x}}^1), \varphi(\bar{\boldsymbol{x}}^2)\}$ 使 $\boldsymbol{a}^0 - \varphi(\bar{\lambda}\bar{\boldsymbol{x}}^1 + (1-\bar{\lambda})\bar{\boldsymbol{x}}^2) \notin K$，即 $\varphi(\bar{\lambda}\bar{\boldsymbol{x}}^1 + (1-\bar{\lambda})\bar{\boldsymbol{x}}^2) \notin \boldsymbol{a}^0 - K$. 据此，由(2.4.22)推知 $\mu_{a^0}(\varphi(\bar{\lambda}\bar{\boldsymbol{x}}^1 + (1-\bar{\lambda})\bar{\boldsymbol{x}}^2), \boldsymbol{\omega}) > 0$，再按(2.4.25)得

$$f_{a^0}(\bar{\lambda}\bar{\boldsymbol{x}}^1 + (1-\bar{\lambda})\bar{\boldsymbol{x}}^2) > 0. \tag{2.4.29}$$

另外，由定义 2.4.3 知 $\boldsymbol{a}^0 \in K - \text{bou}\{\varphi(\bar{\boldsymbol{x}}^1), \varphi(\bar{\boldsymbol{x}}^2)\} \subset (\varphi(\bar{\boldsymbol{x}}^1) + K) \cap (\varphi(\bar{\boldsymbol{x}}^2) + K)$，故有

$$\varphi(\bar{\boldsymbol{x}}^1) \in \boldsymbol{a}^0 - K, \quad \varphi(\bar{\boldsymbol{x}}^2) \in \boldsymbol{a}^0 - K.$$

根据(2.4.22)和(2.4.25)可知

$$f_{a^0}(\bar{x}^1) = \mu_{a^0}\left(\varphi(\bar{x}^1),\ \omega\right) \leqslant 0,$$

$$f_{a^0}(\bar{x}^2) = \mu_{a^0}\left(\varphi(\bar{x}^2),\ \omega\right) \leqslant 0.$$

因此，由(2.4.29)得到

$$f_{a^0}\left(\bar{\lambda}\bar{x}^1 + (1-\bar{\lambda})\bar{x}^2\right) > 0 \geqslant \max\left\{f_{a^0}(\bar{x}^1),\ f_{a^0}(\bar{x}^2)\right\},$$

根据定义 2.3.1 的(1)，导致与对任意的 $a \in \mathscr{Y}$，f_a在 S 上是拟凸的相矛盾.

(2)的证明与(1)的必要性证明部分类似，由定义 2.4.4 的(2)和定义 2.3.1 的(2)，利用引理 2.4.9 和引理 2.4.8 的(2)可以推证. ▯

定理 2.4.12 设 $S \subset \mathscr{X}$ 是非空凸集. 若 $\varphi: S \to \mathscr{Y}$ 是 K-半连续的 K-严格拟凸映射，则 φ 是 S 上的 K-拟凸映射.

证明 对于任意的 $a \in \mathscr{Y}$，因为 φ 在 S 上是 K-半连续的，由引理 2.4.10 知对应的 f_a 是下半连续的. 又根据 φ 在 S 上是 K-严格拟凸的，由引理 2.4.11 的(2)，f_a 在 S 上是严格拟凸的. 现由 f_a 是下半连续的严格拟凸函数，根据定理 2.3.3 得知 f_a 是拟凸函数，因此由引理 2.4.11 的(1)即知 φ 在 S 上是 K-拟凸的. ▯

以下讨论锥拟凸映射的几种条件. 它们都是拟凸函数条件的推广，因而有的文献中也将它们作为锥拟凸映射的定义.

定理 2.4.13 设 $S \subset \mathscr{X}$ 是非空凸集，$\varphi: S \to \mathscr{Y}$ 是映射. 若对任意的 $x^1,\ x^2 \in S$ 和任意的 $\lambda \in (0,\ 1)$，有

$$\varphi(x^1) - \varphi\left(\lambda x^1 + (1-\lambda)x^2\right) \in K$$

或

$$\varphi(x^2) - \varphi\left(\lambda x^1 + (1-\lambda)x^2\right) \in K \tag{2.4.30}$$

(Ferro 的恰当 K-拟凸映射，见文献[22])，则 φ 在 S 上是 K-拟凸的.

证明 不妨设有 $\varphi(x^1) - \varphi(\lambda x^1 + (1-\lambda)x^2) \in K$，则

$$\varphi(\boldsymbol{x}^1) \in \varphi\left(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2\right) + K.$$

据定义 2.4.3 知 $K-\mathrm{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\} \subset \varphi(\boldsymbol{x}^1) + K$，故由上式并注意到 K 是凸锥而得到

$$K-\mathrm{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\} \subset \left(\varphi(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) + K\right) + K$$

$$\subset \varphi\left(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2\right) + K,$$

即

$$K-\mathrm{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\} - \varphi\left(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2\right) \subset K.$$

因此，由定义 2.4.4 的(1)即知 φ 在 S 上是 K-拟凸的. ▯

注 2.4.10　在定理 2.4.13 中的(2.4.30)式中将 K 换作 $\mathrm{int}K$ 或 $K\backslash\{\boldsymbol{0}\}$，并且补加条件 $\varphi(\boldsymbol{x}^1) \neq \varphi(\boldsymbol{x}^2)$，则相应地可以得到 φ 在 S 上是 K-严格拟凸的或 K-次严格拟凸的. 若将(2.4.30)中的 K 换作 $\mathrm{int}K$，并加上条件 $\boldsymbol{x}^1 \neq \boldsymbol{x}^2$，则相应地可得到 φ 在 S 上是 K-强拟凸的. 它们的证明与定理 2.4.13 的证明类似.

定理 2.4.14　设 $S \subset \mathscr{X}$ 是非空凸集. $\varphi: S \to \mathscr{Y}$ 是 K-拟凸映射当且仅当对任意的 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in S$ 和任意的 $\lambda \in (0,1)$，有

$$\bar{\boldsymbol{y}} - \varphi\left(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2\right) \in K$$

$$\forall \bar{\boldsymbol{y}} \in \left\{\boldsymbol{y} \in \mathscr{Y} \mid \boldsymbol{y} - \varphi(\boldsymbol{x}^1) \in K,\ \boldsymbol{y} - \varphi(\boldsymbol{x}^2) \in K\right\} \tag{2.4.31}$$

(Luc 的 K-拟凸映射，见文献[23]).

证明　充分性. 条件(2.4.31)即

$$\bar{\boldsymbol{y}} \in \varphi\left(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2\right) + K$$

$$\forall \bar{\boldsymbol{y}} \in \left(\varphi(\boldsymbol{x}^1) + K\right) \cap \left(\varphi(\boldsymbol{x}^2) + K\right).$$

由定义 2.4.3 知

$$K-\mathrm{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\} \subset \left(\varphi(\boldsymbol{x}^1) + K\right) \cap \left(\varphi(\boldsymbol{x}^2) + K\right),$$

故得 $K - \mathrm{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\} \subset \varphi(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) + K$，或即

$$K - \mathrm{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\} - \varphi(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \subset K.$$

因此，由定义 2.4.4 中的(1)得知 φ 在 S 上是 K-拟凸的.

必要性. 设 φ 在 S 上是 K-拟凸的，则对任意的 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in S$ 和任意的 $\lambda \in (0, 1)$ 有(2.4.14)，也即

$$K - \mathrm{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\} \subset \varphi(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) + K. \tag{2.4.32}$$

任取 $\tilde{\boldsymbol{y}} \in \{\boldsymbol{y} \in \mathscr{Y} : \boldsymbol{y} - \varphi(\boldsymbol{x}^1) \in K, \boldsymbol{y} - \varphi(\boldsymbol{x}^2) \in K\}$，即

$$\tilde{\boldsymbol{y}} \in \{\varphi(\boldsymbol{x}^1) + K\} \cap \{\varphi(\boldsymbol{x}^2) + K\},$$

由定义 2.4.3 可知存在 $\bar{\boldsymbol{y}} \in K - \mathrm{bou}\{\varphi(\boldsymbol{x}^1), \varphi(\boldsymbol{x}^2)\}$ 使 $\tilde{\boldsymbol{y}} \in \bar{\boldsymbol{y}} + K$. 据此，由于(2.4.32)和 K 是凸锥而有

$$\tilde{\boldsymbol{y}} \in (\varphi(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) + K) + K \subset \varphi(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) + K.$$

从而得到 $\tilde{\boldsymbol{y}} - \varphi(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \in K$. ▯

注 2.4.11 将定理 2.4.14 的 $\tilde{\boldsymbol{y}} - \varphi(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) \in K$ 中的 K 分别换作 $\mathrm{int}K$（或 $K\setminus\{\boldsymbol{0}\}$，并加 $\varphi(\boldsymbol{x}^1) \neq \varphi(\boldsymbol{x}^2)$）和 $\mathrm{int}K$（加 $\boldsymbol{x}^1 \neq \boldsymbol{x}^2$），则相应地有 φ 在 S 上是 K-严格拟凸的（或 K-次严格拟凸的）和 K-强拟凸的. 其证明与定理 2.4.14 的证明类似.

设 $\mathscr{Y}^*$ 是 $\mathscr{Y}$ 的对偶空间，$K^* \subset \mathscr{Y}^*$ 是 K 的对偶锥.

定理 2.4.15 设 $S \subset \mathscr{X}$ 是非空凸集，$\varphi: S \to \mathscr{Y}$ 是映射. 若对任意的 $\lambda \in (0, 1)$ 和任意的 $\boldsymbol{\nu}^* \in K^*\setminus\{\boldsymbol{0}\}$ 有

$$\begin{aligned}&\langle \boldsymbol{\nu}^*, \varphi(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2)\rangle \\ &\leqslant \max\{\langle \boldsymbol{\nu}^*, \varphi(\boldsymbol{x}^1)\rangle, \langle \boldsymbol{\nu}^*, \varphi(\boldsymbol{x}^2)\rangle\} \quad \forall \boldsymbol{x}^1, \boldsymbol{x}^2 \in S\end{aligned} \tag{2.4.33}$$

（胡毓达等的 K-拟凸映射，见文献[24]），则 φ 在 S 上是 K-拟

凸的.

证明　用反证法. 假设 φ 在 S 上不是 K-拟凸的，则由定义 2.4.4 的(1)可知，存在 $\bar{x}^1, \bar{x}^2 \in S$ 和 $\bar{\lambda} \in (0, 1)$ 有

$$K - \mathrm{bou}\{\varphi(\bar{x}^1), \varphi(\bar{x}^2)\} - \varphi(\bar{\lambda}\bar{x}^1 + (1-\bar{\lambda})\bar{x}^2) \not\subset K.$$

于是，可取 $y^0 \in K - \mathrm{bou}\{\varphi(\bar{x}^1), \varphi(\bar{x}^2)\}$ 使

$$y^0 - \varphi(\bar{\lambda}\bar{x}^1 + (1-\bar{\lambda})\bar{x}^2) \notin K.$$

由此，单点集 $\{y^0 - \varphi(\bar{\lambda}\bar{x}^1 + (1-\bar{\lambda})\bar{x}^2)\}$ 与闭凸集 K 是严格可分离的，故由定理 1.3.7 的(2)得知存在 $\bar{\nu}^* \in \mathscr{Y}^*\backslash\{\boldsymbol{0}\}$ 使

$$\langle\bar{\nu}^*, y^0 - \varphi(\bar{\lambda}\bar{x}^1 + (1-\bar{\lambda})\bar{x}^2)\rangle < \langle\bar{\nu}^*, y\rangle \quad \forall y \in K.$$

我们指出，实际上 $\bar{\nu}^* \in K^*\backslash\{\boldsymbol{0}\}$(因为否则，若 $\bar{\nu}^* \notin K^*$，由定义 1.4.6 的(1) 知，对某一 $z \in K$ 有 $\langle\bar{\nu}^*, z\rangle < 0$，从而对任意大的 $\lambda > 0$ 有 $\langle\bar{\nu}^*, \lambda z\rangle < 0$. 令 $y = \lambda z \in K$ 代入上式，则得其左端是有限值，而右端可任意小，导致矛盾). 于是，在上式中取 $y = \boldsymbol{0} \in K$，则得存在 $\bar{\nu}^* \in K^*\backslash\{\boldsymbol{0}\}$ 使

$$\langle\bar{\nu}^*, y^0\rangle < \langle\bar{\nu}^*, \varphi(\bar{\lambda}\bar{x}^1 + (1-\bar{\lambda})\bar{x}^2)\rangle. \tag{2.4.34}$$

另外，由定义 2.4.3，有

$$y^0 \in K - \mathrm{bou}\{\varphi(\bar{x}^1), \varphi(\bar{x}^2)\} \subset \{\varphi(\bar{x}^1) + K\} \cap \{\varphi(\bar{x}^2) + K\},$$

故知

$$y^0 - \varphi(\bar{x}^1) \in K, \ y^0 - \varphi(\bar{x}^2) \in K.$$

因为 $\bar{\nu}^* \in K^*\backslash\{\boldsymbol{0}\}$，所以有

$$\langle\bar{\nu}^*, y^0 - \varphi(\bar{x}^1)\rangle \geqslant 0, \quad \langle\bar{\nu}^*, y^0 - \varphi(\bar{x}^2)\rangle \geqslant 0,$$

从而

$$\langle\bar{\nu}^*, y^0\rangle \geqslant \max\{\langle\bar{\nu}^*, \varphi(\bar{x}^1)\rangle, \langle\bar{\nu}^*, \varphi(\bar{x}^2)\rangle\}.$$

利用(2.4.34),得到

$$\langle \bar{\nu}^*, \varphi(\bar{\lambda}\bar{x}^1 + (1-\bar{\lambda})\bar{x}^2) \rangle$$
$$> \max\{\langle \bar{\nu}^*, \varphi(\bar{x}^1)\rangle, \langle \bar{\nu}^*, \varphi(\bar{x}^2)\rangle\},$$

导致与已设条件(2.4.33)矛盾. ▯

注 2.4.12 在定理 2.4.15 的(2.4.33)中将"≤"换作"<",并且补加条件 $\varphi(x^1) \neq \varphi(x^2)$ 或 $x^1 \neq x^2$,则相应地可得 φ 在 S 上是 K-严格拟凸的或 K-强拟凸的. 证明与定理 2.4.15 的证明类似.

最后,介绍锥伪凸映射类.

定义 2.4.5 设 $S \subset \mathscr{X}$ 是非空开凸集,映射 $\varphi: S \to \mathscr{Y}$ 是 Gâteaux 可微的.

(1) 若有

$$\varphi(x^1) - \varphi(x^2) \in \mathrm{int}K \Rightarrow \varphi'_{x^1}(x^2 - x^1) \in -\mathrm{int}K \quad \forall x^1, x^2 \in S, \tag{2.4.35}$$

则称 φ 是集合 S 上的 K-伪凸映射,或映射 φ 在 S 上是 K-伪凸的.

(2) 若有

$$\varphi(x^1) - \varphi(x^2) \in K \Rightarrow \varphi'_{x^1}(x^2 - x^1) \in -\mathrm{int}K$$
$$\forall x^1, x^2 \in S,\ \varphi(x^1) \neq \varphi(x^2), \tag{2.4.36}$$

则称 φ 是集合 S 上的 K-严格伪凸映射,或映射 φ 在 S 上是 K-严格伪凸的.

(3) 若有

$$\varphi(x^1) - \varphi(x^2) \in K \Rightarrow \varphi'_{x^1}(x^2 - x^1) \in -\mathrm{int}K$$
$$\forall x^1, x^2 \in S,\ x^1 \neq x^2, \tag{2.4.37}$$

则称 φ 是集合 S 上的 K-强伪凸映射,或映射 φ 在 S 上是 K-强伪凸的.

(4) 若有

$$\varphi(x^1) - \varphi(x^2) \in K \Rightarrow \varphi'_{x^1}(x^2 - x^1) \in -K \quad \forall x^1, x^2 \in S, \tag{2.4.38}$$

则称 φ 是集合 S 上的 K-次伪凸映射，或映射 φ 在 S 上是 K-次伪凸的.

注 2.4.13　从定义 2.4.5 直接可知，开凸集上的 K-强伪凸映射也是 K-伪凸的，K-严格伪凸的和 K-次伪凸的，又 K-严格伪凸映射也是 K-伪凸的，反之不然.

注 2.4.14　在定义 2.4.5 中，若 $\mathscr{Y}=R$ 和 $K=R_+$，则 $\varphi=f$: $S\to R$ 是实值函数，(2.4.35)和(2.4.36)成为

$$f(x^1)>f(x^2)\Rightarrow f'_{x^1}(x^2-x^1)<0\quad \forall x^1,\ x^2\in S.$$

由此，按定义 2.3.3 的(1)，f 在 S 上是 R_+-伪凸的和 R_+-严格伪凸的，即 f 是 S 上的伪凸函数. 同理，这时(2.4.37)成为(2.3.12)，(2.4.38)成为(2.3.13)，因此 f 在 S 上是 R_+-强(次)伪凸的，即 f 是 S 上的强(次)伪凸函数.

注 2.4.15　在定义 2.4.5 中当 $\mathscr{Y}=R^m$ 和 $K=R^m_+$ ($m\geqslant 2$) 时，$\varphi=f$:$S\to R^m$ 是向量函数. 这时，不难得知：f 是 R^m_+-伪凸映射即它的每一分量函数都是伪凸函数，f 是 R^m_+-强伪凸映射即它的每一分量函数都是强伪凸函数，f 是 R^m_+-次伪凸映射即它的每一分量函数都是次伪凸函数.

定理 2.4.16　设 $S\subset\mathscr{X}$ 是非空开凸集，映射 φ: $S\to\mathscr{Y}$ 是 Gâteaux 可微的.

(1) 若 φ 是 S 上的 K-凸映射，则 φ 是 S 上的 K-伪凸映射和 K-次伪凸映射.

(2) 若 φ 是 S 上的 K-严格凸映射，则 φ 是 S 上的 K-严格伪凸映射.

(3) 若 φ 是 S 上的 K-强凸映射，则 φ 是 S 上的 K-强伪凸映射.

证明　(1) 由 φ 在 S 上是 K-凸的，根据定义 2.4.1 的(1)可知，对任意的 x^1, $x^2\in S$ 和任意的 $\lambda\in(0,1)$ 有

$$\lambda\varphi(x^2)+(1-\lambda)\varphi(x^1)-\varphi\big(\lambda x^2+(1-\lambda)x^1\big)\in K,$$

也即

$$\varphi(x^1+\lambda(x^2-x^1))-\varphi(x^1)\in-\lambda[\varphi(x^1)-\varphi(x^2)]-K.$$

现设 $\varphi(x^1)-\varphi(x^2)\in \text{int}K$，因为 K 是凸锥，故有

$$\frac{\varphi(x^1+\lambda(x^2-x^1))-\varphi(x^1)}{\lambda}\in-[\varphi(x^1)-\varphi(x^2)]-K$$

$$\subset-\text{int}K-K=-\text{int}K.$$

令 $\lambda\to 0$，按定义 2.3.2，得

$$\varphi'_{x^1}(x^2-x^1)=\lim_{\lambda\to 0}\frac{\varphi(x^1+\lambda(x^2-x^1))-\varphi(x^1)}{\lambda}\in-\text{int}K.$$

因此，由定义 2.4.5 中的(1)即知 φ 在 S 上是 K-伪凸的. 若设 $\varphi(x^1)-\varphi(x^2)\in K$，同理可推得 $\varphi'_{x^1}(x^2-x^1)\in-K$，故由定义 2.4.5 的(4)知 φ 在 S 上是 K-次伪凸的.

(2)和(3)与(1)的证明类似. □

定理 2.4.17 设 $S\subset\mathscr{X}$ 是非空开凸集，映射 $\varphi: S\to\mathscr{Y}$ 是 Gâteaux 可微的. 若 φ 是 S 上的 K-拟凸映射，则 φ 是 S 上的 K-次伪凸映射.

证明 由 φ 在 S 上是 K-拟凸的，按定义 2.4.4 中的(1)可知，对任意的 $x^1, x^2\in S$ 和任意的 $\lambda\in(0,1)$ 有

$$K-\text{bou}\{\varphi(x^2),\varphi(x^1)\}-\varphi(\lambda x^2+(1-\lambda)x^1)\subset K. \tag{2.4.39}$$

设 $\varphi(x^1)-\varphi(x^2)\in K$，则 $\varphi(x^1)\in\varphi(x^2)+K$，从而

$$(\varphi(x^1)+K)\cap(\varphi(x^2)+K)=\varphi(x^1)+K.$$

于是，由定义 2.4.3 可得 $K-\text{bou}\{\varphi(x^2),\varphi(x^1)\}=\varphi(x^1)$. 据此，由(2.4.39)知 $\varphi(x^1)-\varphi(\lambda x^2+(1-\lambda)x^1)\in K$，从而有

$$\frac{\varphi(x^1+\lambda(x^2-x^1))-\varphi(x^1)}{\lambda}\in-K.$$

令 $\lambda\to 0$,由定义 2.3.2 得

$$\varphi'_{x^1}(x^2-x^1)=\lim_{\lambda\to 0}\frac{\varphi\big(x^1+\lambda(x^2-x^1)\big)-\varphi(x^1)}{\lambda}\in -K.$$

因此,由定义 2.4.5 的(4),即知 φ 在 S 上是 K-次伪凸的. ▯

定理 2.4.18 设 $S\subset\mathscr{X}$ 是非空开凸集,映射 $\varphi: S\to\mathscr{Y}$ 是 Gâteaux 可微的. 若对任意的 $\lambda\in(0,1)$ 和任意的 $\nu^*\in K^*\setminus\{\boldsymbol{0}\}$,有

$$\langle\nu^*,\varphi(x^1)\rangle>\langle\nu^*,\varphi(x^2)\rangle\Rightarrow$$

$$\langle\nu^*,\varphi'_{x^1}(x^2-x^1)\rangle<0\quad\forall x^1,x^2\in S,\tag{2.4.40}$$

则 φ 在 S 上是 K-伪凸的.

证明 用反证法,假设在已设条件下 φ 在 S 上不是 K-伪凸的,则由定义 2.4.5 的(1)可知,存在 $\bar{x}^1,\bar{x}^2\in S$ 和 $\bar{\lambda}\in(0,1)$,有

$$\varphi(\bar{x}^1)-\varphi(\bar{x}^2)\in\operatorname{int}K\Rightarrow\varphi'_{\bar{x}^1}(\bar{x}^2-\bar{x}^1)\notin-\operatorname{int}K.\tag{2.4.41}$$

上式右端表明单点集 $\{-\varphi'_{\bar{x}^1}(\bar{x}^2-\bar{x}^1)\}$ 与开凸集 $\operatorname{int}K$ 可以分离,于是由定理 1.3.7 的(1)得知,存在 $\nu^*\in\mathscr{Y}^*\setminus\{\boldsymbol{0}\}$ 使

$$\langle\nu^*,-\varphi'_{\bar{x}^1}(\bar{x}^2-\bar{x}^1)\rangle\leqslant\langle\nu^*,y\rangle\quad\forall y\in\operatorname{int}K.\tag{2.4.42}$$

与定理 2.4.15 的证明中同理有 $\nu^*\in K^*\setminus\{\boldsymbol{0}\}$,将它与(2.4.41)的左端相乘,由定理 1.4.11 的(1)知

$$\langle\nu^*,\varphi(\bar{x}^1)-\varphi(\bar{x}^2)\rangle>0,$$

或

$$\langle\nu^*,\varphi(\bar{x}^1)\rangle>\langle\nu^*,\varphi(\bar{x}^2)\rangle.\tag{2.4.43}$$

另外,对于(2.4.42),令 $y\to 0$,得到

$$\langle\nu^*,\varphi'_{\bar{x}^1}(\bar{x}^2-\bar{x}^1)\rangle\geqslant 0.\tag{2.4.44}$$

(2.4.43)和(2.4.44)与(2.4.40)相矛盾.□

注 2.4.16　在定理 2.4.18 的(2.4.40)中,将左端的“＞”换作“⩾”并补加条件 $\varphi(\boldsymbol{x}^1)\neq\varphi(\boldsymbol{x}^2)$ 或 $\boldsymbol{x}^1\neq\boldsymbol{x}^2$,相应地可以得到 φ 在 S 上是 K-严格伪凸的或 K-强伪凸的.若将(2.4.40)中的“＞”和“＜”分别换作“⩾”和“⩽”,则可得到 φ 在 S 上是 K-次伪凸的.它们的证明与定理 2.4.18 的证明类似.

第 3 章　凸函数的次微分

凸函数的次微分是凸分析的核心内容，同时也是非光滑分析的重要组成部分.

这一章研究凸函数在一点处的变化率或微分问题. 可以看到，在通常的意义下，凸函数不一定是可微的. 为此，本章引进更能刻画凸函数在一点处变化特征的次梯度和次微分概念，并且讨论它们的有关性质. 最后，介绍次微分在凸规划研究中的应用.

§3.1　凸函数的方向导数

本节先考察凸函数沿给定方向的变化率也即方向导数问题. 可以看到，凸函数变化的一个直观特征是：它在其有效域中的一点处沿任何单边方向的方向导数总是存在的. 但一般地说，它在同一点处沿不同方向的方向导数却不一定是相同的.

设 $\mathscr{X}$ 是线性拓扑空间.

定义 3.1.1　设 $f: \mathscr{X} \to R \cup \{\pm\infty\}$ 是实值函数，点 $\boldsymbol{x}^0 \in \operatorname{dom} f$，$\boldsymbol{d} \in \mathscr{X}$.

(1) 若极限

$$f'_+(\boldsymbol{x}^0; \boldsymbol{d}) = \lim_{t \to 0^+} \frac{f(\boldsymbol{x}^0 + t\boldsymbol{d}) - f(\boldsymbol{x}^0)}{t} \qquad (3.1.1)$$

存在，则称它是函数 f 在点 $\boldsymbol{x}^0$ 处沿方向 $\boldsymbol{d}$ 的(单边)右方向导数. 若极限

$$f'_-(\boldsymbol{x}^0; \boldsymbol{d}) = \lim_{t \to 0^-} \frac{f(\boldsymbol{x}^0 + t\boldsymbol{d}) - f(\boldsymbol{x}^0)}{t} \qquad (3.1.2)$$

存在，则称它是函数 f 在点 $\boldsymbol{x}^0$ 处沿方向 $\boldsymbol{d}$ 的(单边)左方向导数.

(2) 若 $f'_+(\boldsymbol{x}^0;\boldsymbol{d}) = f'_-(\boldsymbol{x}^0;\boldsymbol{d})$，即极限

$$f'(\boldsymbol{x}^0;\boldsymbol{d}) = \lim_{t\to 0}\frac{f(\boldsymbol{x}^0+t\boldsymbol{d})-f(\boldsymbol{x}^0)}{t} \tag{3.1.3}$$

存在，则称它是函数 f 在点 $\boldsymbol{x}^0$ 处沿方向 $\boldsymbol{d}$ 的方向导数.

注 3.1.1 由定义 3.1.1 的(3)和定义 2.3.2 可知，若 f 在点 $\boldsymbol{x}^0$ 处沿任意方向 $\boldsymbol{d}\in\mathscr{H}$ 的方向导数均存在并且相等，则 f 在点 $\boldsymbol{x}^0$ 处是 Gâteaux 可微的，映射 $f'_{x^0}:\mathscr{H}\to R$, $\boldsymbol{d}\mapsto f'_{x^0}(\boldsymbol{d}) = f'(\boldsymbol{x}^0;\boldsymbol{d})$ 即是 f 在点 $\boldsymbol{x}^0$ 处的 Gâteaux 导数.

定理 3.1.1 设 $f:\mathscr{H}\to R\cup\{+\infty\}$ 是正常凸函数，$\boldsymbol{x}\in \operatorname{int}(\operatorname{dom}f)$，则对任意的 $\boldsymbol{d}\in\mathscr{H}$，$f'_+(\boldsymbol{x};\boldsymbol{d})$ 和 $f'_-(\boldsymbol{x};\boldsymbol{d})$ 存在，并且

$$f'_-(\boldsymbol{x};\boldsymbol{d})\leqslant f'_+(\boldsymbol{x};\boldsymbol{d}).$$

证明 设 $t\in R$，记 $g(t)=f(\boldsymbol{x}+t\boldsymbol{d})$. 因为 f 是正常凸的，由定理 2.1.8 和定义 2.1.5 可知，$g(t)$ 是 R 上的正常凸函数，并且 $g(0)=f(\boldsymbol{x})$ 为有限值. 据此，可以推知 $\dfrac{g(t)-g(0)}{t}$ 是 $t>0$ 的非减函数，从而极限 $\lim\limits_{t\to 0^+}\dfrac{g(t)-g(0)}{t}$ 存在. 由定义 3.1.1 的(1)即得

$$\begin{aligned} f'_+(\boldsymbol{x};\boldsymbol{d}) &= \lim_{t\to 0^+}\frac{f(\boldsymbol{x}+t\boldsymbol{d})-f(\boldsymbol{x})}{t} \\ &= \lim_{t\to 0^-}\frac{g(t)-g(0)}{t} \end{aligned}$$

存在. 同理，可以推知 $f'_-(\boldsymbol{x};\boldsymbol{d})$ 存在，并且易知有 $f'_-(\boldsymbol{x};\boldsymbol{d})\leqslant f'_+(\boldsymbol{x};\boldsymbol{d})$. □

定理 3.1.1 指出正常凸函数在其有效域内的任一点处沿任何方向的单边方向导数总是存在的，并且其左方向导数值不超过右方向导数值. 但是，下面的例子说明，凸函数在其有效域中的一点处其方向导数可能不存在，因而由注 3.1.1 可知，它在该点处不是

Gâteaux 可微的.

例 3.1.1 范数函数的方向导数. 考虑例 2.1.1 中的范数函数

$$f(\boldsymbol{x}) = \|\boldsymbol{x}\|, \boldsymbol{x} \in R^n.$$

在例 2.1.1 中已指出它是 R^n 上的凸函数. 考察点 $\boldsymbol{0} \in R^n$ 和任一单位向量 $\boldsymbol{d} \in R^n$, $\|\boldsymbol{d}\| = 1$. 由(3.1.1)和(3.1.2)可知, f 在点 $\boldsymbol{0}$ 处沿 $\boldsymbol{d}$ 的(单边)右方向导数和左方向导数分别是

$$f'_+(\boldsymbol{0};\boldsymbol{d}) = \lim_{t \to 0^+} \frac{\|\boldsymbol{0} + t\boldsymbol{d}\| - \|\boldsymbol{0}\|}{t} = 1$$

和

$$f'_-(\boldsymbol{0};\boldsymbol{d}) = \lim_{t \to 0^-} \frac{\|\boldsymbol{0} + t\boldsymbol{d}\| - \|\boldsymbol{0}\|}{t} = -1.$$

由于 $f'_+(\boldsymbol{0};\boldsymbol{d}) \neq f'_-(\boldsymbol{0};\boldsymbol{d})$, 故按定义 3.1.1 的(2), f 在点 $\boldsymbol{0}$ 处沿任何方向 $\boldsymbol{d}$ 的方向导数均不存在.

凸函数的单边方向导数有以下性质.

定理 3.1.2 设 $f: \mathscr{X} \to R \cup \{+\infty\}$ 是正常凸函数, $\boldsymbol{x} \in \mathrm{dom} f$, $\boldsymbol{d} \in \mathscr{X}$, 则 $f'_+(\boldsymbol{x};\boldsymbol{d})$ 关于 $\boldsymbol{d}$ 是正常凸函数.

证明 对于任意的 $\boldsymbol{d}^1, \boldsymbol{d}^2 \in \mathscr{X}$ 和任意的 $\lambda \in (0, 1)$, 由 f 是凸函数, 我们有

$$\begin{aligned}
& f(\boldsymbol{x} + t(\lambda \boldsymbol{d}^1 + (1-\lambda)\boldsymbol{d}^2)) - f(\boldsymbol{x}) \\
= & f(\lambda(\boldsymbol{x} + t\boldsymbol{d}^1) + (1-\lambda)(\boldsymbol{x} + t\boldsymbol{d}^2)) - f(\boldsymbol{x}) \\
\leqslant & \lambda f(\boldsymbol{x} + t\boldsymbol{d}^1) + (1-\lambda) f(\boldsymbol{x} + t\boldsymbol{d}^2) - f(\boldsymbol{x}) \\
= & \lambda[f(\boldsymbol{x} + t\boldsymbol{d}^1) - f(\boldsymbol{x})] + (1-\lambda)[f(\boldsymbol{x} + t\boldsymbol{d}^2) - f(\boldsymbol{x})].
\end{aligned}$$

两边除以 t, 并令 $t \to 0^+$, 得到

$$f'_+(\boldsymbol{x}; \lambda \boldsymbol{d}^1 + (1-\lambda)\boldsymbol{d}^2)$$

$$= \lim_{t \to 0^+} \frac{f(\boldsymbol{x} + t(\lambda \boldsymbol{d}^1 + (1-\lambda)\boldsymbol{d}^2)) - f(\boldsymbol{x})}{t}$$

$$\leqslant \lambda \lim_{t \to 0^+} \frac{f(\boldsymbol{x} + t\boldsymbol{d}^1) - f(\boldsymbol{x})}{t}$$

$$+ (1-\lambda) \lim_{t \to 0^+} \frac{f(\boldsymbol{x} + t\boldsymbol{d}^2) - f(\boldsymbol{x})}{t}$$

$$= \lambda f'_+(x; \boldsymbol{d}^1) + (1-\lambda) f'_+(x; \boldsymbol{d}^2).$$

据此，由定义 2.1.1 的(1)即知 $f'_+(\boldsymbol{x}; \boldsymbol{d})$关于 $\boldsymbol{d}$ 是凸函数. 又由 f 是正常凸函数，按定义 3.1.1 的(1)，进而可知 $f'_+(\boldsymbol{x}; \boldsymbol{d})$关于 $\boldsymbol{d}$ 是正常凸的. □

定理 3.1.3 设 $f: \mathscr{X} \to R \cup \{+\infty\}$ 是正常凸函数，$\boldsymbol{d} \in \mathscr{X}$，$f$ 在点 $\boldsymbol{x} \in \mathrm{dom} f$ 处是连续的.

(1) $f'_+(\boldsymbol{x}; \boldsymbol{d})$关于 $\boldsymbol{d}$ 是连续的.

(2) $f'_+(\boldsymbol{x}; \boldsymbol{d})$关于 $\boldsymbol{d}$ 是正齐次和次可加的(即次线性的).

证明 (1) 由定理 3.1.2 得知 $f'_+(\boldsymbol{x}; \boldsymbol{d})$关于 $\boldsymbol{d}$ 是正常凸函数. 因此，为证明 $f'_+(\boldsymbol{x}; \boldsymbol{d})$关于 $\boldsymbol{d}$ 是连续的，依据定理 2.2.1，只需证明它在某邻域内上有界. 事实上，从 f 在点 $\boldsymbol{x}$ 处是连续的可知，对于某零点邻域中的 $\boldsymbol{d}$，$f(\boldsymbol{x} + \boldsymbol{d})$ 上有界. 于是，由

$$f'_+(\boldsymbol{x}; \boldsymbol{d}) = \lim_{t \to 0^+} \frac{f(\boldsymbol{x} + t\boldsymbol{d}) - f(\boldsymbol{x})}{t} \leqslant f(\boldsymbol{x} + \boldsymbol{d}) - f(\boldsymbol{x})$$

即知 $f'_+(\boldsymbol{x}; \boldsymbol{d})$在此零点邻域内也上有界.

(2) 对于 $\alpha \geqslant 0$，由(3.1.1)有

$$f'_+(\boldsymbol{x}; \alpha\boldsymbol{d}) = \lim_{t \to 0^+} \frac{f(\boldsymbol{x} + t\alpha\boldsymbol{d}) - f(\boldsymbol{x})}{t}$$

$$= \alpha \lim_{t \to 0^+} \frac{f(\boldsymbol{x} + t\alpha\boldsymbol{d}) - f(\boldsymbol{x})}{\alpha t}$$

$$= \alpha f'_+(\boldsymbol{x}; \boldsymbol{d}),$$

故正齐次性成立. 又对任意的 $\boldsymbol{d}^1, \boldsymbol{d}^2 \in \mathscr{X}$，因 f 是凸函数，故

$$
\begin{aligned}
f'_+(\boldsymbol{x};\boldsymbol{d}^1+\boldsymbol{d}^2) &= \lim_{t\to 0^+}\frac{f(\boldsymbol{x}+t(\boldsymbol{d}^1+\boldsymbol{d}^2))-f(\boldsymbol{x})}{t}\\
&\leqslant \lim_{t\to 0^+}\frac{f(\boldsymbol{x}+2t\boldsymbol{d}^1)+f(\boldsymbol{x}+2t\boldsymbol{d}^2)-2f(\boldsymbol{x})}{2t}\\
&= \lim_{t\to 0^+}\frac{f(\boldsymbol{x}+2t\boldsymbol{d}^1)-f(\boldsymbol{x})}{2t}+\lim_{t\to 0^+}\frac{f(\boldsymbol{x}+2t\boldsymbol{d}^2)-f(\boldsymbol{x})}{2t}\\
&= f'_+(\boldsymbol{x};\boldsymbol{d}^1)+f'_+(\boldsymbol{x};\boldsymbol{d}^2).
\end{aligned}
$$

由此，$f'_+(\boldsymbol{x};\boldsymbol{d})$关于 $\boldsymbol{d}$ 具有次可加性. 综合以上结果，$f'_+(\boldsymbol{x};\boldsymbol{d})$关于 $\boldsymbol{d}$ 是次线性泛函. □

对于 Gâteaux 可微的正常凸函数，有如下结论.

设 $\mathscr{H}^*$ 是 $\mathscr{H}$ 的对偶空间.

定理 3.1.4　设 $f:\mathscr{H}\to R\cup\{+\infty\}$ 是正常凸函数，$\boldsymbol{x}\in \mathrm{dom}f$. 若 f 在点 $\boldsymbol{x}$ 处是 Gâteaux 可微的，则 $f'_x(\boldsymbol{d})=f'(\boldsymbol{x};\boldsymbol{d})$ 关于 $\boldsymbol{d}$ 是连续线性泛函，即存在连续线性泛函 $\boldsymbol{x}^*\in\mathscr{H}^*$，使

$$
f'_x(\boldsymbol{d})=\langle \boldsymbol{x}^*,\boldsymbol{d}\rangle\quad \forall \boldsymbol{d}\in\mathscr{H}. \tag{3.1.4}
$$

证明　因为 f 在点 $\boldsymbol{x}$ 处是 Gâteaux 可微的，由定理 2.3.7，则 f 在点 $\boldsymbol{x}$ 处是连续的. 因此，根据定理 3.1.3 得知 $f'_+(\boldsymbol{x};\boldsymbol{d})$关于 $\boldsymbol{d}$ 是连续的次线性泛函.

又从 f 在点 $\boldsymbol{x}$ 处是 Gâteaux 可微的，由定义 2.3.2 和定义 3.1.1 的(2)可知，对任意的 $\boldsymbol{d}\in\mathscr{H}$，有 $f'_x(\boldsymbol{d})=f'(\boldsymbol{x};\boldsymbol{d})=f'_+(\boldsymbol{x};\boldsymbol{d})=-f'_+(\boldsymbol{x};-\boldsymbol{d})$. 于是，由已知 $f'_x(\boldsymbol{d})=f'_+(\boldsymbol{x};\boldsymbol{d})$ 和 $-f'_x(\boldsymbol{d})=f'_+(\boldsymbol{x};-\boldsymbol{d})$ 关于 $\boldsymbol{d}$ 都是连续的次线性泛函，即得$f'_x(\boldsymbol{d})$关于 $\boldsymbol{d}$ 是连续线性泛函. □

注 3.1.2　从定理 3.1.4 得知，对于在某点 $\boldsymbol{x}\in\mathscr{H}$ 处 Gâteaux 可微的正常凸函数，它在该点处的 Gâteaux 导数 f'_x 与满足

$$
f'(\boldsymbol{x};\boldsymbol{d})=\langle \boldsymbol{x}^*,\boldsymbol{d}\rangle\quad \forall \boldsymbol{d}\in\mathscr{H}
$$

的连续线性泛函 $\boldsymbol{x}^*$ 对应. 因而也可将此连续线性泛函定义为它在

该点处的 Gâteaux 导数.

注 3.1.3　在微积分中我们知道，对于 Euclid 空间上的可微函数 $f: R^n \to R$，设 $\nabla f(\boldsymbol{x})$ 是它在点 $\boldsymbol{x} \in R^n$ 处的梯度，$f'(\boldsymbol{x}; \boldsymbol{d})$ 是在点 $\boldsymbol{x}$ 处沿任一方向 $\boldsymbol{d}$ 的方向导数，则有 $f'(\boldsymbol{x}; \boldsymbol{d}) = \nabla f(\boldsymbol{x})^T \boldsymbol{d}$，或即

$$f'_x(\boldsymbol{d}) = \langle \nabla f(\boldsymbol{x}), \boldsymbol{d} \rangle \quad \forall \boldsymbol{d} \in R^n.$$

将它与(3.1.4)当 $\mathscr{X} = R^n$ 时相比，即知当 f 是正常凸函数时，与它在点 x 处的 Gâteaux 导数对应的连续线性泛函即是梯度 $\nabla f(\boldsymbol{x})$. 可见，Gâteaux 导数是梯度的一种推广.

下面介绍另一可微的概念.

设 $\mathscr{B}$ 是 Banach 空间，$\mathscr{B}^*$ 是 $\mathscr{B}$ 的对偶空间.

定义 3.1.2　设 $f: \mathscr{B} \to R \cup \{\pm\infty\}$ 是广义实值函数，点 $\boldsymbol{x}^0 \in \mathrm{dom} f$，$\boldsymbol{y} \in \mathscr{B}$. 若存在 $\boldsymbol{x}' \in \mathscr{B}^*$ 使得

$$\lim_{\boldsymbol{y} \to \boldsymbol{x}^0} \frac{f(\boldsymbol{y}) - f(\boldsymbol{x}^0) - \langle \boldsymbol{x}', \boldsymbol{y} - \boldsymbol{x}^0 \rangle}{\| \boldsymbol{y} - \boldsymbol{x}^0 \|} = 0, \qquad (3.1.5)$$

则称函数 f 在点 $\boldsymbol{x}^0$ 处是 Frechet 可微的. 称 $\boldsymbol{x}'$ 是 f 在点 $\boldsymbol{x}^0$ 处的 Frechet 导数，记作 $f'(\boldsymbol{x}^0) = \boldsymbol{x}'$. 并且称 $\langle \boldsymbol{x}', \boldsymbol{y} - \boldsymbol{x}^0 \rangle$ 是 f 在点 $\boldsymbol{x}^0$ 处的 Frechet 微分，记作 $\mathrm{d}f(\boldsymbol{x}^0) = \langle \boldsymbol{x}', \boldsymbol{y} - \boldsymbol{x}^0 \rangle$.

注 3.1.4　显然，(3.1.5)与

$$f(\boldsymbol{y}) - f(\boldsymbol{x}^0) = \langle \boldsymbol{x}', \boldsymbol{y} - \boldsymbol{x}^0 \rangle + o(\| \boldsymbol{y} - \boldsymbol{x}^0 \|), \ \boldsymbol{y} \in \mathscr{B} \qquad (3.1.6)$$

等价，其中 $\dfrac{o(\| \boldsymbol{y} - \boldsymbol{x}^0 \|)}{\| \boldsymbol{y} - \boldsymbol{x}^0 \|} \to 0 (\boldsymbol{y} \to \boldsymbol{x}^0)$. 对于 Euclid 空间上的可微函数 $f: R^n \to R$，设 $\mathrm{d}f(\boldsymbol{x}^0)$ 是它在点 $\boldsymbol{x}^0 \in R^n$ 处的全微分，则有

$$f(\boldsymbol{y}) - f(\boldsymbol{x}^0) = \mathrm{d}f(\boldsymbol{x}^0) + o(\| \boldsymbol{y} - \boldsymbol{x}^0 \|).$$

当 $\mathscr{B} = R^n$ 时将它与(3.1.6)相比，即知这时 Frechet 微分即是全微分：

$$\langle x', y-x^0\rangle = \mathrm{d}f(x^0).$$

可见，Frechet 微分是全微分的推广，我们仍记它为 $\mathrm{d}f(x^0)$.

定理 3.1.5　设 $f: \mathscr{B} \to R \cup \{\pm\infty\}$ 是广义实值函数，$x \in \mathrm{dom}f$. 若 f 在点 x 处是 Frechet 可微的，则 f 在点 x 处是 Gâteaux 可微的，并且

$$f'_x(d) = \langle f'(x), d\rangle \quad \forall d \in \mathscr{B}.$$

证明　对任何 $d \in \mathscr{B}$，令 $y = x + td$（其中 $t \in R$）. 由定义 3.1.1 的(2)和(3.1.5)有

$$\begin{aligned} f'(x; d) &= \lim_{t\to 0} \frac{f(x+td) - f(x)}{t} \\ &= \lim_{y\to x} \frac{f(y) - f(x) - \langle f'(x), y-x\rangle}{\| y-x \|} \| d \| \\ &\quad + \langle f'(x), d\rangle \\ &= \langle f'(x), d\rangle. \end{aligned}$$

因为 $d \in \mathscr{B}$ 是任意的，所以由定义 2.3.2 和定义 3.1.1 的(2)即得 $f'_x(d) = \langle f'(x), d\rangle$. ▯

推论 3.1.6　设 $f: \mathscr{B} \to R \cup \{+\infty\}$ 是正常凸函数，$x \in \mathrm{dom}f$. 若 f 在点 x 处是 Frechet 可微的，则 $f'(x)$是 $\mathscr{B}$ 上的连续线性泛函.

证明　由定理 3.1.5 和定理 3.1.4 可得. ▯

§ 3.2　次梯度和次微分

从上一节我们知道，凸函数在其有效域内虽然总存在单边方向导数，但却不一定是 Gâteaux 可微的. 本节引入凸函数的次微分概念，并且阐明在次微分的意义下，凸函数在其连续点处总是次可微的.

设 $\mathscr{X}$ 是线性拓扑空间，$\mathscr{X}^*$ 是 $\mathscr{X}$ 的对偶空间.

定义 3.2.1 设 $f: \mathscr{X} \to R \cup \{\pm\infty\}$ 是凸函数，点 $\boldsymbol{x}^0 \in \mathscr{X}$. 若 $\boldsymbol{x}^* \in \mathscr{X}^*$，并且

$$\langle \boldsymbol{x}^*, \boldsymbol{y} - \boldsymbol{x}^0 \rangle \leqslant f(\boldsymbol{y}) - f(\boldsymbol{x}^0) \quad \forall \boldsymbol{y} \in \mathscr{X}, \tag{3.2.1}$$

则称 $\boldsymbol{x}^*$ 是 f 在点 $\boldsymbol{x}^0$ 处的*次梯度*. f 在点 $\boldsymbol{x}^0$ 处的所有次梯度的集合称为 f 在点 $\boldsymbol{x}^0$ 处的*次微分*，记作 $\partial f(\boldsymbol{x}^0)$，即

$$\partial f(\boldsymbol{x}^0) = \{\boldsymbol{x}^* \in \mathscr{X}^* \mid \langle \boldsymbol{x}^*, \boldsymbol{y} - \boldsymbol{x}^0 \rangle \leqslant f(\boldsymbol{y}) - f(\boldsymbol{x}^0) \quad \forall \boldsymbol{y} \in \mathscr{X}\}. \tag{3.2.2}$$

若 $\partial f(\boldsymbol{x}^0) \neq \varnothing$，则称 f 在点 $\boldsymbol{x}^0$ 处是*次可微*的. 若 f 在集合 $S \subset \mathscr{X}$ 的每一点处是次可微的，则称 f 在 S 上是*次可微*的.

例 3.2.1 范数函数的次微分. 再考虑例 2.1.1 中的凸函数

$$f(\boldsymbol{x}) = \|\boldsymbol{x}\|, \ \boldsymbol{x} \in R^n.$$

由例 3.1.1 已知，它在点 $\boldsymbol{0}$ 处沿任何方向的方向导数均不存在. 但据定义 3.2.1，$\boldsymbol{x}^* \in \mathscr{X}^* = R^n$ 是 f 在点 $\boldsymbol{0}$ 处的次梯度，意即对任意的 $\boldsymbol{y} \in R^n$ 有 $\langle \boldsymbol{x}^*, \boldsymbol{y} \rangle \leqslant \|\boldsymbol{y}\| - \|\boldsymbol{0}\|$，也即 $\boldsymbol{x}^{*T}\boldsymbol{y} \leqslant \|\boldsymbol{y}\|$ 或 $\|\boldsymbol{x}^*\| \leqslant 1$. 由此可知，$R^n$ 中满足 $\|\boldsymbol{x}^*\| \leqslant 1$ 的向量 $\boldsymbol{x}^*$ 都是 f 在点 $\boldsymbol{0}$ 处的次梯度，而 f 在点 $\boldsymbol{0}$ 处的次微分即单位球

$$\partial f(\boldsymbol{0}) = \{\boldsymbol{x}^* \in R^n \mid \|\boldsymbol{x}^*\| \leqslant 1\},$$

因而 f 在点 $\boldsymbol{0}$ 处是次可微的. 特别地，对于 $n = 1$ 时，凸函数 $f(x) = |x|$ 在点 $\boldsymbol{0}$ 处的次微分是闭区间 $\partial f(\boldsymbol{0}) = [-1, 1]$，其任一次梯度 $x^* \in [-1, 1]$（图 3.2.1 中的 x^* 对应 $\zeta = |x|$ 在点 $\boldsymbol{0}$ 处切线的斜率）.

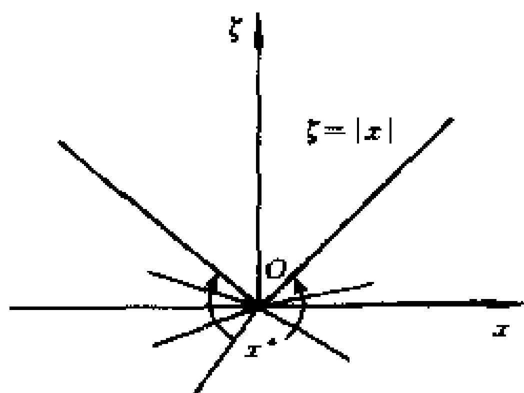


图 3.2.1

注 3.2.1 设 $\boldsymbol{x}^*$ 是 f 在点 $\boldsymbol{x}^0 \in \mathscr{X}$ 处的次梯度. 令 $\beta = \langle \boldsymbol{x}^*, \boldsymbol{x}^0 \rangle - f(\boldsymbol{x}^0)$，考虑 $\mathscr{X}$

$\times R$ 中非垂直的闭超平面

$$\widetilde{H}=\{(\boldsymbol{x},\zeta)\in\mathscr{X}\times R \mid \langle\boldsymbol{x}^*,\boldsymbol{x}\rangle-\zeta=\beta\}, \quad (3.2.3)$$

则有 $(\boldsymbol{x}^0, f(\boldsymbol{x}^0))\in\widetilde{H}$. 由于 f 是凸函数,根据定理 2.1.5 知它的上图象 $\mathrm{epi}f\subset\mathscr{X}\times R$ 是凸集,并且显然有 $(\boldsymbol{x}^0, f(\boldsymbol{x}^0))\in\mathrm{epi}f$. 另外,任取 $(\boldsymbol{y},\eta)\in\mathrm{epi}f$, 由定义 2.1.6 的(1)和定义 3.2.1 得

$$\eta\geqslant f(\boldsymbol{y})\geqslant\langle\boldsymbol{x}^*,\boldsymbol{y}-\boldsymbol{x}^0\rangle+f(\boldsymbol{x}^0),$$

即

$$\langle\boldsymbol{x}^*,\boldsymbol{y}\rangle-\eta\leqslant\langle\boldsymbol{x}^*,\boldsymbol{x}^0\rangle-f(\boldsymbol{x}^0),$$

于是

$$(\boldsymbol{y},\eta)\in\widetilde{H}_-^{\leq}=\{(\boldsymbol{x},\zeta)\in\mathscr{X}\times R \mid \langle\boldsymbol{x}^*,\boldsymbol{x}\rangle-\zeta\leqslant\beta\}.$$

因为 $(\boldsymbol{y},\eta)\in\mathrm{epi}f$ 是任取的,故得 $\mathrm{epi}f\subset\widetilde{H}_-^{\leq}$. 根据以上的讨论,由定义 1.3.3 得知 $\widetilde{H}$ 是 $\mathrm{epi}f$ 在点 $(\boldsymbol{x}^0, f(\boldsymbol{x}^0))$ 处的非垂直的闭支撑超平面(图 3.2.2(a)). 若 f 在点 $\boldsymbol{x}^0$ 处的次梯度不唯一,则 f 的上图象 $\mathrm{epi}f$ 在点 $(\boldsymbol{x}^0, f(\boldsymbol{x}^0))$ 处的支撑超平面也不唯一(图 3.2.2 (b)). 特别是,由(3.2.3)当 $\mathscr{X}=R$ 时,可知单变量凸函数 f 在点 x^0 处的次梯度 x^* 就是它在点 x^0 处支撑切线 $\widetilde{H}$ 的斜率.

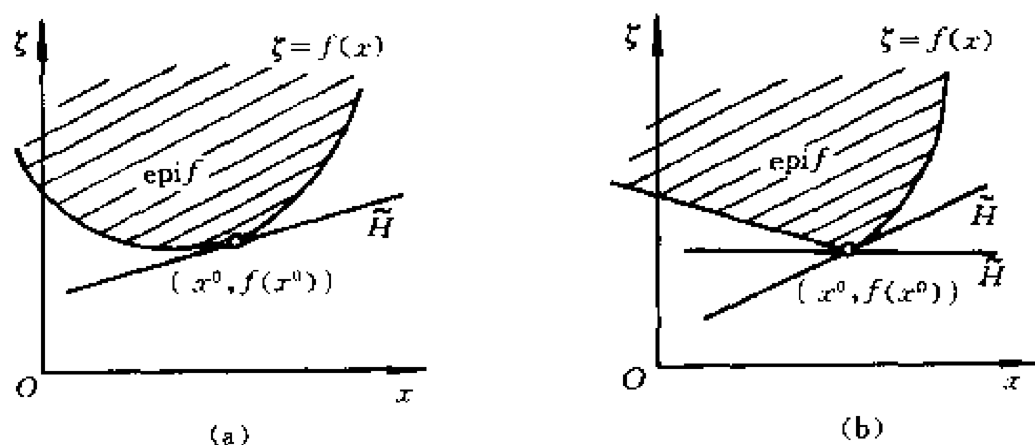


图 3.2.2

注 3.2.2 我们指出,凸函数 f 在任意点 $\boldsymbol{x}\in\mathscr{X}$ 处的次微分 $\partial f(\boldsymbol{x})$ 都是闭凸集. 事实上,若 $\partial f(\boldsymbol{x})$ 是空集或单点集,它显然是

闭凸的. 设 $\partial f(\boldsymbol{x})$ 至少包含两个点. 由(3.2.2)易知它是闭集. 此外,任取 $\boldsymbol{x}_1^*,\ \boldsymbol{x}_2^* \in \partial f(\boldsymbol{x})$ 和任意的 $\lambda \in (0, 1)$, 由(3.2.2)有

$$\langle \boldsymbol{x}_1^*,\ \boldsymbol{y}-\boldsymbol{x}\rangle \leqslant f(\boldsymbol{y}) - f(\boldsymbol{x}) \quad \forall \boldsymbol{y} \in \mathscr{X}, \tag{3.2.4}$$

$$\langle \boldsymbol{x}_2^*,\ \boldsymbol{y}-\boldsymbol{x}\rangle \leqslant f(\boldsymbol{y}) - f(\boldsymbol{x}) \quad \forall \boldsymbol{y} \in \mathscr{X}. \tag{3.2.5}$$

作 $\lambda(3.2.4) + (1-\lambda)(3.2.5)$, 则得

$$\begin{aligned}
&\left\langle \lambda\boldsymbol{x}_1^* + (1-\lambda)\boldsymbol{x}_2^*,\ \boldsymbol{y}-\boldsymbol{x} \right\rangle \\
&\leqslant \left[\lambda f(\boldsymbol{y}) + (1-\lambda)f(\boldsymbol{y})\right] - \left[\lambda f(\boldsymbol{x}) + (1-\lambda)f(\boldsymbol{x})\right] \\
&= f(\boldsymbol{y}) - f(\boldsymbol{x}).
\end{aligned}$$

于是,再由(3.2.2)即知 $\lambda\boldsymbol{x}_1^* + (1-\lambda)\boldsymbol{x}_2^* \in \partial f(\boldsymbol{x})$, 从而 $\partial f(\boldsymbol{x})$ 是凸集. 此外, $\partial f(\boldsymbol{x})$ 还是弱 * 紧的(见[25]定理 4.1.2 的推论 2).

下面是次微分的存在性定理.

定理 3.2.1 设 $f: \mathscr{X} \to R \cup \{+\infty\}$ 是正常凸函数. 若 f 在某点 $\boldsymbol{x}^0 \in \mathrm{int}(\mathrm{dom} f)$ 处是连续的,则

$$\partial f(\boldsymbol{x}) \neq \varnothing \quad \forall \boldsymbol{x} \in \mathrm{int}(\mathrm{dom} f).$$

证明 首先,因为 f 是凸函数,故它的上图象 $\mathrm{epi} f \subset \mathscr{X} \times R$ 是凸集. 其次,由 f 在点 $\boldsymbol{x}^0$ 处是连续的,可知存在 $\boldsymbol{x}^0$ 的邻域 $U(\boldsymbol{x}^0)$ 和 $t > 0$ 使得

$$f(\boldsymbol{x}) \leqslant t \quad \forall \boldsymbol{x} \in U(\boldsymbol{x}^0).$$

因此,当 $\boldsymbol{x} \in U(\boldsymbol{x}^0)$ 和 $\eta \geqslant t$ 时,有 $(\boldsymbol{x}, \eta) \in \mathrm{epi} f$, 从而 $\mathrm{int}(\mathrm{epi} f) \neq \varnothing$. 另外,显然有 $(\boldsymbol{x}^0, f(\boldsymbol{x}^0)) \in \mathrm{rb}(\mathrm{epi} f)$, 于是根据推论 1.3.12 得知存在闭超平面 $H \subset \mathscr{X} \times R$, 它在点 $(\boldsymbol{x}^0, f(\boldsymbol{x}^0))$ 处正常支撑 $\mathrm{epi} f$. 由于 $\boldsymbol{x}^0 \in \mathrm{int}(\mathrm{dom} f)$, 则可推知 H 是非垂直的(见[15]的引理 5.33). 现设

$$H = \{(\boldsymbol{x}, \zeta) \in \mathscr{X} \times R \mid \langle \boldsymbol{x}^*, \boldsymbol{x}\rangle + \alpha\zeta = \beta\},$$

并且不失一般性让 $\mathrm{epi} f \subset H_+^c$, 则对任意的 $(\boldsymbol{y}, \eta) \in \mathrm{epi} f$ 有

$$\langle x^*, y\rangle + \alpha\eta \geqslant \beta.$$

因为 η 可以充分大，故 $\alpha > 0$. 由于特别是对 $y \in \mathrm{dom} f$ 有 $(y, f(y)) \in \mathrm{epi} f$，从上式可知

$$\langle x^*, y\rangle + \alpha f(y) \geqslant \beta. \tag{3.2.6}$$

另外，由 $(x^0, f(x^0))$ 是支撑点，则 $(x^0, f(x^0)) \in H$，因而有

$$\langle x^*, x^0\rangle + \alpha f(x^0) = \beta.$$

将它代入(3.2.6)有 $\langle x^*, y\rangle + \alpha f(y) \geqslant \langle x^*, x^0\rangle + \alpha f(x^0)$，于是得到

$$\left\langle -\frac{1}{\alpha}x^*, y - x^0\right\rangle \leqslant f(y) - f(x^0).$$

由于当 $y \notin \mathrm{dom} f$ 时，按定义 2.1.4 知上式显然也成立，故由定义 3.2.1 得知 $-\dfrac{1}{\alpha}x^* \in \partial f(x^0)$，从而 $\partial f(x^0) \neq \varnothing$.

由于 f 在点 $x^0 \in \mathrm{int}(\mathrm{dom} f)$ 处是连续的，它在 $\mathrm{int}(\mathrm{dom} f)$ 内一点的某邻域内上有界，故根据定理 2.2.1，f 在 $\mathrm{int}(\mathrm{dom} f)$ 内是连续的. 由此，按照上面的证明得知，对任意的 $x \in \mathrm{int}(\mathrm{dom} f)$ 有 $\partial f(x) \neq \varnothing$. □

现在给出次梯度与方向导数之间的关系.

定理 3.2.2 设 $f: \mathscr{H} \to R \cup \{+\infty\}$ 是正常凸函数. 若 f 在点 $x \in \mathrm{dom} f$ 处是连续的，则 $x^* \in \partial f(x)$ 当且仅当

$$\langle x^*, d\rangle \leqslant f'_+(x; d) \quad \forall d \in \mathscr{H}. \tag{3.2.7}$$

证明 必要性. 设 $x^* \in \partial f(x)$，由定义 3.2.1 可知，对任意的 $d \in \mathscr{H}$ 和 $t > 0$ 有

$$\langle x^*, td\rangle \leqslant f(x + td) - f(x).$$

因此，按定义 3.1.1 的(1)即得

$$\langle x^*, d\rangle \leqslant \lim_{t\to 0^+} \frac{f(x+td)-f(x)}{t} = f'_+(x;d).$$

充分性. 对于任意的 $d \in \mathscr{X}$, 易知 $\frac{f(x+td)-f(x)}{t}$ 是 $t>0$ 的非减函数. 据此, 由(3.2.7)得知有

$$\langle x^*, d\rangle \leqslant f'_+(x;d) = \lim_{t\to 0^+} \frac{f(x+td)-f(x)}{t} \leqslant \frac{f(x+td)-f(x)}{t},$$

从而 $\langle x^*, td\rangle \leqslant f(x+td)-f(x)$. 令 $y = x + td$, 则 $\langle x^*, y-x\rangle \leqslant f(y)-f(x)$, 于是由定义 3.2.1 即知 $x^* \in \partial f(x)$. ▯

注 3.2.3　事实上, 在定理 3.2.2 的基础上, 还可以得到

$$f'_+(x;d) = \sup_{x^*\in\partial f(x)} \langle x^*, d\rangle \quad \forall d \in \mathscr{X}$$

(见[25]的定理 4.1.2).

以下定理表明, 在一定意义的可微条件下, 次微分即为该意义下的导数或梯度.

定理 3.2.3　设 $f: \mathscr{X} \to R \cup \{+\infty\}$ 是正常凸函数, f 在点 $x \in \mathrm{dom} f$ 处是连续的. 若 f 在点 x 处是 Gâteaux 可微的, 并且其 Gâteaux 导数是 $f'_x \in \mathscr{X}^*$, 则 $\partial f(x) = \{f'_x\}$.

证明　因为 f 在点 x 处是 Gâteaux 可微的, 故由定理 3.1.4 和注 3.1.2 有

$$f'_x(d) = \langle f'_x, d\rangle \quad \forall d \in \mathscr{X}. \tag{3.2.8}$$

由于按定义 2.3.2, 这时对任意的 $d \in \mathscr{X}$ 有 $f'_x(d) = f'(x;d)$, 因而也有

$$f'_x(d) = f'_+(x;d) \quad \forall d \in \mathscr{X}. \tag{3.2.9}$$

结合(3.2.8), 知

$$\langle f'_x, d\rangle = f'_+(x;d) \quad \forall d \in \mathscr{X},$$

因此由定理 3.2.2 得到 $f'_x \in \partial f(x)$.

假设另有 $x^* \in \partial f(x)$，则由定理 3.2.2 和(3.2.9)有

$$\langle x^*, d \rangle \leqslant f'_x(d) \quad \forall d \in \mathscr{H}.$$

据此，注意(3.2.8)得

$$\langle f'_x - x^*, d \rangle \geqslant 0 \quad \forall d \in \mathscr{H}.$$

由 d 的任意性便导致 $x^* = f'_x$. ▯

设 $\mathscr{B}$ 是 Banach 空间.

定理 3.2.4 设 $f: \mathscr{B} \to R \cup \{+\infty\}$ 是正常凸函数，f 在点 $x \in \mathrm{dom} f$ 处是连续的. 若 f 在点 x 处是 Frechet 可微的，并且其 Frechet 导数是 $f'(x)$，则 $\partial f(x) = \{f'(x)\}$.

证明 由 f 在点 x 处是 Frechet 可微的，据定理 3.1.5，它在该点处也是 Gâteaux 可微的，并且

$$f'_x(d) = \langle f'(x), d \rangle \quad \forall d \in \mathscr{B}.$$

设 $f'_x \in \mathscr{H}^*$ 是 f 在点 x 处的 Gâteaux 导数，由定理 3.1.4 和注 3.1.2 对比(3.1.4)和上式，得知 $f'(x) = f'_x$. 由定理 3.2.3 即得 $\partial f(x) = \{f'(x)\}$. ▯

推论 3.2.5 设 $f: R^n \to R$ 是正常凸函数. 若 f 在点 $x \in \mathrm{dom} f$ 处是可微的，则 $\partial f(x) = \{\nabla f(x)\}$.

证明 由注 3.1.3 和定理 3.2.3 可得证. ▯

例 3.2.2 仿射函数的次微分. 设 $a \in \mathscr{H}$, $\beta \in R$, 从例2.1.1已知仿射函数

$$l(x) = \langle a, x \rangle + \beta, \quad x \in \mathscr{H}$$

是 $\mathscr{H}$ 上的凸函数. 由定义 3.2.1，$x^* \in \mathscr{H}^*$ 是函数 l 在点 $x \in \mathscr{H}$ 处的次梯度当且仅当对任意的 $y \in \mathscr{H}$ 有

$$\langle x^*, y - x \rangle \leqslant \langle a, y \rangle + \beta - \langle a, x \rangle - \beta,$$

或即 $\langle a - x^*, y - x \rangle \geqslant 0$. 由 y 的任意性，即知 $x^* = a$，于是

$\partial l(\boldsymbol{x}) = \{\boldsymbol{a}\}$.

例 3.2.3 指示函数的次微分. 设 $S \subset \mathscr{X}$ 是非空凸集，考虑例 2.1.1 中的指示函数

$$\delta_S(\boldsymbol{x}) = \begin{cases} 0, & \boldsymbol{x} \in S; \\ +\infty, & \boldsymbol{x} \in \mathscr{X} \setminus S. \end{cases}$$

例 2.1.1 已指出它是 S 上的凸函数. 由定义 3.2.1，它在点 $\boldsymbol{x} \in S$ 处的次微分是

$$\partial \delta_S(\boldsymbol{x}) = \{x^* \in \mathscr{X}^* \mid \langle \boldsymbol{x}^*, \boldsymbol{y} - \boldsymbol{x} \rangle \leqslant \delta_S(\boldsymbol{y}) - \delta_S(\boldsymbol{x}) \quad \forall \boldsymbol{y} \in \mathscr{X}\},$$

即

$$\partial \delta_S(\boldsymbol{x}) = \{\boldsymbol{x}^* \in \mathscr{X}^* \mid \langle \boldsymbol{x}^*, \boldsymbol{y} - \boldsymbol{x} \rangle \leqslant 0 \quad \forall \boldsymbol{y} \in S\}. \tag{3.2.10}$$

由注 3.2.2 知，对于任何 $\boldsymbol{x} \in S$，(3.2.10)右端的集合都是闭凸的. 此外，对任意的 $\lambda > 0$，显然有 $\langle \lambda \boldsymbol{x}^*, \boldsymbol{y} - \boldsymbol{x} \rangle \leqslant 0 (\forall \boldsymbol{y} \in S)$，因而从(3.2.10)由定义 1.4.1 和定义 1.4.2 还得知 $\partial \delta_S(\boldsymbol{x})$是 $\mathscr{X}^*$ 中的闭凸锥. 特别是，当 $S = K$ 是凸锥时，由(3.2.10)有

$$\partial \delta_S(\boldsymbol{0}) = \{\boldsymbol{x}^* \in \mathscr{X}^* \mid \langle \boldsymbol{x}^*, \boldsymbol{y} \rangle \leqslant 0 \quad \forall \boldsymbol{y} \in K\}.$$

按定义 1.4.6 中的(3)，它是凸锥 K 的负对偶锥，即

$$K^- = \{x^* \in \mathscr{X}^* \mid \langle \boldsymbol{x}^*, \boldsymbol{y} \rangle \leqslant 0 \quad \forall \boldsymbol{y} \in K\} = \partial \delta_S(\boldsymbol{0}). \tag{3.2.11}$$

利用指示函数的次微分，我们给出凸集的法锥和切锥概念如下.

定义 3.2.2 设 $S \subset \mathscr{X}$ 是非空凸集，点 $\boldsymbol{x}^0 \in S$.

(1) 集合

$$N_S(\boldsymbol{x}^0) = \{x^* \in \mathscr{X}^* \mid \langle \boldsymbol{x}^*, \boldsymbol{y} - \boldsymbol{x}^0 \rangle \leqslant 0 \quad \forall \boldsymbol{y} \in S\} \tag{3.2.12}$$

称为凸集 S 在点 x^0 处的法锥，$x^* \in N_S(x^0)$ 称为 S 在点 x^0 处的法向量.

(2) 集合

$$T_S(x^0) = \{d \in \mathscr{H} \mid \langle x^*, d\rangle \leqslant 0 \quad \forall x^* \in N_S(x^0)\} \tag{3.2.13}$$

称为凸集 S 在点 x^0 处的切锥，$d \in T_S(x^0)$ 称为 S 在点 x^0 处的切向量.

注 3.2.4　设 $x \in S$，由定义 3.2.2 不难验证 $N_S(x)$ 是 $\mathscr{H}^*$ 中的锥. 对比(3.2.12)和(3.2.10)可知，凸集 S 在点 x 处的法锥即是 S 上的指示函数在该点处的次微分：$N_S(x) = \partial\delta_S(x)$. 根据注 3.2.2，$\partial\delta_S(x)$ 是闭凸的，故 $N_S(x)$ 是闭凸锥. 由(3.2.13)和(3.2.11)又得知，凸集 S 在点 x 处的切锥即是该点处的法锥的负对偶锥：$T_S(x) = N_S(x)^-$.

注 3.2.5　在定义 3.2.2 中当 $\mathscr{H} = R^n$ 并把对偶积 $\langle\cdot,\cdot\rangle$ 理解为数量积时，由(3.2.12)可知 $N_S(x^0)$ 就是与所有向量 $y - x^0(y \in S)$ 的夹角不小于 $\pi/2$ 的向量全体，而 $T_S(x^0)$ 则是把原点移到 x^0 处后由 S 所生成的闭凸锥(图 3.2.3). 若 S 在点 x^0 处是"光滑"的，则 $T_S(x^0)$ 是过原点的半空间.

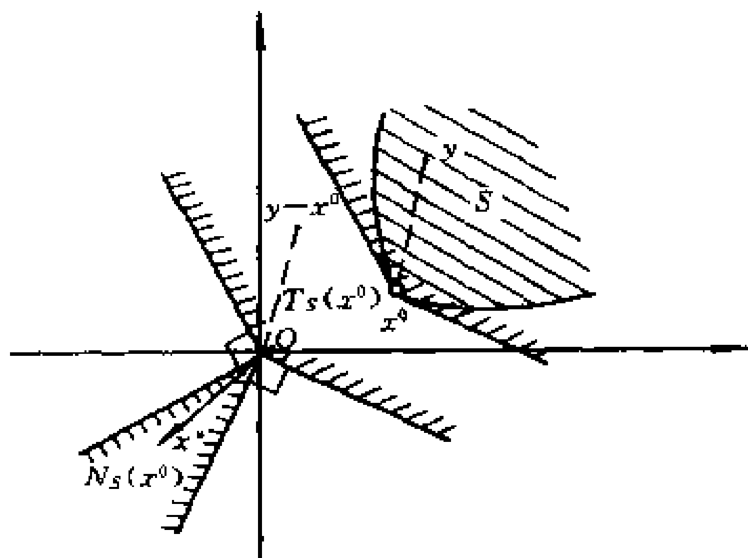


图 3.2.3

最后，介绍凸函数的次微分与共轭函数之间的关系.

定理 3.2.6 设 $f: \mathscr{X} \to R \cup \{+\infty\}$ 是正常凸函数，$f^*: \mathscr{X}^* \to R \cup \{\pm\infty\}$ 是 f 的共轭函数，$\boldsymbol{x} \in \mathrm{dom} f$，则 $\boldsymbol{x}^* \in \partial f(\boldsymbol{x})$ 当且仅当

$$\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle = f(\boldsymbol{x}) + f^*(\boldsymbol{x}^*). \tag{3.2.14}$$

证明 由定义 3.2.1，$\boldsymbol{x}^* \in \partial f(\boldsymbol{x})$ 意味着

$$\langle \boldsymbol{x}^*, \boldsymbol{y} - \boldsymbol{x} \rangle \leqslant f(\boldsymbol{y}) - f(\boldsymbol{x}) \quad \forall \boldsymbol{y} \in \mathscr{X},$$

即

$$\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle - f(\boldsymbol{x}) \geqslant \langle \boldsymbol{x}^*, \boldsymbol{y} \rangle - f(\boldsymbol{y}) \quad \forall \boldsymbol{y} \in \mathscr{X}.$$

再由定义 2.2.4，上式等价于

$$\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle - f(\boldsymbol{x}) = \sup_{\boldsymbol{y} \in \mathscr{X}} \{\langle \boldsymbol{x}^*, \boldsymbol{y} \rangle - f(\boldsymbol{y})\} = f^*(\boldsymbol{x}^*),$$

此即(3.2.14). ▯

推论 3.2.7 设 $f: \mathscr{X} \to R \cup \{+\infty\}$ 是正常凸函数，f^* 和 f^{**} 分别是 f 的共轭函数和二次共轭函数，$\boldsymbol{x} \in \mathrm{dom} f$，并且 $f(\boldsymbol{x}) = f^{**}(\boldsymbol{x})$.

(1) $\boldsymbol{x}^* \in \partial f(\boldsymbol{x})$ 当且仅当 $\boldsymbol{x} \in \partial f^*(\boldsymbol{x}^*)$.

(2) $\partial f(\boldsymbol{x}) = \partial f^{**}(\boldsymbol{x})$.

证明 (1) 由定理 3.2.6 和 $f(\boldsymbol{x}) = f^{**}(\boldsymbol{x})$，则 $\boldsymbol{x}^* \in \partial f(\boldsymbol{x})$ 等价于

$$\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle = f^{**}(\boldsymbol{x}) + f^*(\boldsymbol{x}^*) = f^*(\boldsymbol{x}^*) + (f^*)^*(\boldsymbol{x}).$$

再利用定理 3.2.6 得知，它又等价于 $\boldsymbol{x} \in \partial f^*(\boldsymbol{x}^*)$.

(2) 由 $f(\boldsymbol{x}) = f^{**}(\boldsymbol{x})$ 知 $f(\boldsymbol{x}^*) = (f^{**})^*(\boldsymbol{x}^*)$. 于是，由定理 3.2.6 得知 $\boldsymbol{x}^* \in \partial f(\boldsymbol{x})$ 等价于

$$\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle = f(\boldsymbol{x}) + f^*(\boldsymbol{x}^*) = f^{**}(\boldsymbol{x}) + (f^{**})^*(\boldsymbol{x}^*).$$

再由定理 3.2.6 即知，它又等价于 $\boldsymbol{x}^* \in \partial f^{**}(\boldsymbol{x})$. ▯

推论 3.2.8 设 $f: \mathscr{X} \to R \cup \{+\infty\}$ 是正常凸函数，f^{**} 是

f 的二次共轭函数，$x \in \mathrm{dom} f$，并且 $\partial f(x) \neq \varnothing$.

(1) $f(x) = f^{**}(x)$.

(2) $f(x) = \mathrm{cl} f(x)$.

证明 (1) 由 $\partial f(x) \neq \varnothing$，设 $x^* \in \partial f(x)$，据定理 3.2.6 有

$$\langle x^*, x \rangle = f(x) + f^*(x^*).$$

另外，对 f^* 利用(2.2.10)(Young-Fenchel 不等式)，则有

$$\langle x^*, x \rangle \leqslant f^{**}(x) + f^*(x^*).$$

由以上两式得到 $f(x) \leqslant f^{**}(x)$. 因为根据定理 2.2.12 的(1)知 f^{**} 是 f 的弱函数，即有 $f^{**}(x) \leqslant f(x)$，所以 $f(x) = f^{**}(x)$.

(2) 因 f^{**} 是 f 的弱函数，故也有 $f^{**}(x) \leqslant \mathrm{cl} f(x) \leqslant f(x)$. 于是，由(1)即知 $f(x) = \mathrm{cl} f(x)$. ▯

§ 3.3 次微分的性质

设 $\mathscr{X}$ 是线性拓扑空间. 由定义 3.2.1 我们知道，凸函数 f 在点 $x \in \mathscr{X}$ 处的次微分 $\partial f(x)$ 是对偶空间 $\mathscr{X}^*$ 中的一个集合，因而它确定一个从 $\mathscr{X}$ 到 $\mathscr{X}^*$ 中集合的集值映射 $\partial f: \mathscr{X} \to 2^{\mathscr{X}^*}$，$x \mapsto \partial f(x)$. 我们称此集值映射为 f 的次微分映射，也简称次微分. 本节阐述次微分映射的运算性质、连续性和单调性.

先介绍次微分的简单运算性质.

定理 3.3.1 设 $f, f_1, f_2: \mathscr{X} \to R \cup \{+\infty\}$ 是正常凸函数，$x \in \mathscr{X}$.

(1) 若 $\alpha \geqslant 0$，则 $\partial(\alpha f)(x) = \alpha \partial f(x)$.

(2) $\partial f_1(x) + \partial f_2(x) \subset \partial(f_1 + f_2)(x)$.

证明 (1) 当 $\alpha = 0$ 时，是显然的. 设 $\alpha > 0$，由定义 3.2.1 得知 $x^* \in \partial(\alpha f)(x)$ 意味着对任意的 $y \in \mathscr{X}$ 有

$$\langle x^*, y - x \rangle \leqslant \alpha f(y) - \alpha f(x),$$

或

$$\left\langle \frac{x^*}{\alpha},\ y - x \right\rangle \leqslant f(y) - f(x).$$

再由定义 3.2.1 有 $\frac{x^*}{\alpha} \in \partial f(x)$，即 $x^* \in \alpha\partial f(x)$.

(2) 设 $x^* \in \partial f_1(x) + \partial f_2(x)$，则存在 x_1^* 和 x_2^* 使 $x^* = x_1^* + x_2^*$，并且 $x_1^* \in \partial f_1(x)$ 和 $x_2^* \in \partial f_2(x)$. 由定义 3.2.1 可知

$$\langle x_1^*,\ y - x\rangle \leqslant f_1(y) - f_1(x) \quad \forall y \in \mathscr{H},$$

$$\langle x_2^*,\ y - x\rangle \leqslant f_2(y) - f_2(x) \quad \forall y \in \mathscr{H}.$$

将以上两式的两端相加，得

$$\langle x_1^* + x_2^*,\ y - x\rangle \leqslant [f_1(y) + f_2(y)] - [f_1(x) + f_2(x)] \quad \forall y \in \mathscr{H},$$

再按定义 3.2.1，于是得 $x^* = x_1^* + x_2^* \in \partial(f_1 + f_2)(x)$. □

定理 3.3.2 设 $f: \mathscr{H} \to R \cup \{+\infty\}$ 是正常凸函数，$a \in \mathscr{H}^*$，$\beta \in R$，$l(x) = \langle a,\ x\rangle + \beta$ 是 $\mathscr{H}$ 上的仿射函数，则

$$\partial(f + l)(x) = \partial f(x) + \partial l(x) \quad \forall x \in \mathscr{H}.$$

证明 由定义 3.2.1 可知，$x^* \in \partial(f + l)(x)$ 意味着

$$\langle x^*,\ y - x\rangle \leqslant [f(y) + \langle a,\ y\rangle + \beta] - [f(x) + \langle a,\ x\rangle + \beta] \quad \forall y \in \mathscr{H},$$

或即

$$\langle x^* - a,\ y - x\rangle \leqslant f(y) - f(x) \quad \forall y \in \mathscr{H}.$$

于是，再按定义 3.2.1 有 $x^* - a \in \partial f(x)$，由例 3.2.2 知 $\partial l(x) = \{a\}$，因而得 $x^* \in \partial f(x) + \partial l(x)$. □

定理 3.3.3 设 $f_i: \mathscr{H} \to R \cup \{+\infty\}(i = 1, \cdots, m)$ 是正常凸函数，

$$f(x) = \max_{1 \leqslant i \leqslant m} f_i(x),\ x \in \mathscr{H},$$

$I(\boldsymbol{x}) = \{i \in \{1, \cdots, m\} \mid f_i(\boldsymbol{x}) = f(\boldsymbol{x}),\ \boldsymbol{x} \in \mathscr{X}\}$. 若各 f_i 在点 $\boldsymbol{x}^0 \in \mathrm{dom} f$ 处是连续的，则

$$\partial f(\boldsymbol{x}^0) = \mathrm{co} \bigcup_{i \in I(\boldsymbol{x}^0)} \partial f_i(\boldsymbol{x}^0).$$

证明　因为各 $f_i (i = 1, \cdots, m)$ 在 $\boldsymbol{x}^0$ 处是连续的，所以由定理 3.2.2 有

$$\partial f_i(\boldsymbol{x}^0) = \{\boldsymbol{x}^* \in \mathscr{X}^* \mid \langle \boldsymbol{x}^*, \boldsymbol{d} \rangle \leqslant f'_{i+}(\boldsymbol{x}^0; \boldsymbol{d}) \quad \forall \boldsymbol{d} \in \mathscr{X}\},$$

$$i = 1, \cdots, m. \tag{3.3.1}$$

又由 f_i 在 $\boldsymbol{x}^0$ 处是连续的得知，在 $\boldsymbol{x}^0$ 的某邻域内有

$$f(\boldsymbol{x}) = \max_{i \in I(\boldsymbol{x}^0)} f_i(\boldsymbol{x}), \tag{3.3.2}$$

并且它在该点处是连续的. 于是同样由定理 3.2.2 又有

$$\partial f(\boldsymbol{x}^0) = \{\boldsymbol{x}^* \in \mathscr{X}^* \mid \langle \boldsymbol{x}^*, \boldsymbol{d} \rangle \leqslant f'_+(\boldsymbol{x}^0; \boldsymbol{d}) \quad \forall \boldsymbol{d} \in \mathscr{X}\}. \tag{3.3.3}$$

现在证明

$$f'_+(\boldsymbol{x}^0; \boldsymbol{d}) = \max_{i \in I(\boldsymbol{x}^0)} f'_{i+}(\boldsymbol{x}^0; \boldsymbol{d}), \quad \boldsymbol{d} \in \mathscr{X}. \tag{3.3.4}$$

事实上，当 $i \in I(\boldsymbol{x}^0)$ 时，由定义 3.1.1 的(1)和(3.3.2)有

$$\begin{aligned} f'_{i+}(\boldsymbol{x}^0; \boldsymbol{d}) &= \lim_{t \to 0^+} \frac{f_i(\boldsymbol{x}^0 + t\boldsymbol{d}) - f_i(\boldsymbol{x}^0)}{t} \\ &\leqslant \lim_{t \to 0^+} \frac{f(\boldsymbol{x}^0 + t\boldsymbol{d}) - f(\boldsymbol{x}^0)}{t} \\ &= f'_+(\boldsymbol{x}^0; \boldsymbol{d}), \end{aligned}$$

因此

$$f'_+(\boldsymbol{x}^0; \boldsymbol{d}) \geqslant \max_{i \in I(\boldsymbol{x}^0)} f'_{i+}(\boldsymbol{x}^0; \boldsymbol{d}). \tag{3.3.5}$$

另一方面，从定义 3.1.1 的(1)可知有正数列 $\{t_k\}$，$t_k \to 0$ $(k \to$

∞)，并且

$$f'_+(x^0;\ d) = \lim_{k\to\infty} \frac{f(x^0 + t_k d) - f(x^0)}{t_k}. \qquad (3.3.6)$$

由(3.3.2)得知，至少存在一个 $\bar{i} \in I(x^0)$ 使对无限多个 k 有

$$f(x^0 + t_k d) = f_{\bar{i}}(x^0 + t_k d).$$

由于对这个 $\bar{i} \in I(x^0)$ 根据(3.3.6)有

$$f'_+(x^0;\ d) = \lim_{k\to\infty} \frac{f_{\bar{i}}(x^0 + t_k d) - f_{\bar{i}}(x^0)}{t_k} = f'_{\bar{i}+}(x^0;\ d),$$

故由(3.3.5)知(3.3.4)成立.

根据(3.3.4)和注 3.2.3 可知，对 $d \in \mathscr{X}$ 有

$$f'_+(x^0;\ d) = \max_{i\in I(x^0)} f'_{i+}(x^0;\ d)$$

$$= \max_{i\in I(x^0)} \sup_{x^*\in\partial f_i(x^0)} \langle x^*,\ d\rangle,$$

即

$$f'_+(x^0;\ d) = \sup_{x^*\in \bigcup\limits_{i\in I(x^0)} \partial f_i(x^0)} \langle x^*,\ d\rangle,\ d \in \mathscr{X}.$$

由此，按定义 1.3.4 得知 $f'_+(x^0;\ d)$关于 d 是集合 $\bigcup\limits_{i\in I(x^0)} \partial f_i(x^0)$ 的支撑函数. 因为显然有 $ri \bigcup\limits_{i\in I(x^0)} \partial f_i(x^0) \neq \varnothing$，所以由定理 1.3.13 和(3.3.3)得到

$$\overline{\mathrm{co}} \bigcup_{i\in I(x^0)} \partial f_i(x^0) = \{x^* \in \mathscr{X}^* \mid \langle x^*,\ d\rangle \leqslant f'_+(x^0;\ d)$$

$$\forall d \in \mathscr{X}\} = \partial f(x^0).$$

最后，证明 $\mathrm{co} \bigcup\limits_{i\in I(x^0)} \partial f_i(x^0)$ 是闭集，从上式即可得定理的结论. 为此，作函数 $g: R^m \times (\mathscr{X}^*)^m \to R$,

$$g(\boldsymbol{\lambda},\ \boldsymbol{x}_1^*,\ \cdots,\ \boldsymbol{x}_m^*) = \sum_{i=1}^{m}\lambda_i \boldsymbol{x}_i^*, \tag{3.3.7}$$

其中 $\boldsymbol{\lambda} = (\lambda_1,\ \cdots,\ \lambda_m)^T \in R^m$. 记

$$\Lambda = \left\{\boldsymbol{\lambda} = (\lambda_1,\ \cdots,\ \lambda_m)^T \in R^m \,|\, \lambda_i \geqslant 0\right.$$

$$\left.(i = 1,\ \cdots,\ m),\ \sum_{i=1}^{m}\lambda_i = 1\right\},$$

由(3.3.7)可以推知

$$g\left(\Lambda \times \left(\prod_{i\in I(\boldsymbol{x}^0)} \partial f_i(\boldsymbol{x}^0)\right)\right) = \mathrm{co} \bigcup_{i\in I(\boldsymbol{x}^0)} \partial f_i(\boldsymbol{x}^0). \tag{3.3.8}$$

因为 Λ 和 $\partial f_i(\boldsymbol{x}^0)(i \in I(x^0))$ 都是闭集，并且由(3.3.7)知 g 关于 $\boldsymbol{\lambda}$ 和 $\boldsymbol{x}_i^*(i = 1,\ \cdots,\ m)$ 是连续的，所以根据(3.3.8)即得 $\mathrm{co} \bigcup_{i\in I(\boldsymbol{x}^0)} \partial f_i(\boldsymbol{x}^0)$ 是闭集. □

下面是次微分运算的基本定理.

定理 3.3.4　设 $f_1, f_2: \mathscr{H} \to R \cup \{+\infty\}$ 是正常凸函数. 若存在点 $\boldsymbol{x}^0 \in (\mathrm{dom} f_1) \cap (\mathrm{dom} f_2)$，使得 f_1 在点 $\boldsymbol{x}^0$ 处是连续的，则

$$\partial(f_1 + f_2)(\boldsymbol{x}) = \partial f_1(\boldsymbol{x}) + \partial f_2(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \mathscr{H}. \tag{3.3.9}$$

证明　由定理 3.3.1 的(2)，只需再证明

$$\partial(f_1 + f_2)(\boldsymbol{x}) \subset \partial f_1(\boldsymbol{x}) + \partial f_2(\boldsymbol{x}).$$

为此，设 $\boldsymbol{x}^* \in \partial(f_1 + f_2)(\boldsymbol{x})$，我们证明有 $\boldsymbol{x}^* \in \partial f_1(\boldsymbol{x}) + \partial f_2(\boldsymbol{x})$. 不妨设 $\boldsymbol{x}^* = \boldsymbol{0}$，$\boldsymbol{x} = \boldsymbol{0}$，$f_1(\boldsymbol{0}) + f_2(\boldsymbol{0}) = 0$，因为否则，可令

$$g_1(z) = f_1(\boldsymbol{x} + z) - f_1(\boldsymbol{x}) - \langle \boldsymbol{x}^*,\ z\rangle,$$

$$g_2(\boldsymbol{z}) = f_2(\boldsymbol{x} + \boldsymbol{z}) - f_2(\boldsymbol{x}),$$

将问题转为对 g_1 和 g_2 来讨论.

以下设 $\boldsymbol{0}\in\partial(f_1+f_2)(\boldsymbol{0})$，我们来证明 $\boldsymbol{0}\in\partial f_1(\boldsymbol{0})+\partial f_2(\boldsymbol{0})$. 从 $\boldsymbol{0}\in\partial(f_1+f_2)(\boldsymbol{0})$，由定义 3.2.1 知

$$\langle\boldsymbol{0},\boldsymbol{y}-\boldsymbol{0}\rangle\leqslant(f_1+f_2)(\boldsymbol{y})-(f_1+f_2)(\boldsymbol{0})\quad\forall\boldsymbol{y}\in\mathscr{H},$$

从而有

$$\inf_{\boldsymbol{y}\in\mathscr{H}}\left[f_1(\boldsymbol{y})+f_2(\boldsymbol{y})\right]=f_1(\boldsymbol{0})+f_2(\boldsymbol{0})=0.\tag{3.3.10}$$

作集合

$$S_1=\operatorname{epi}f_1=\left\{(\boldsymbol{x},\eta)\in\mathscr{H}\times R\,|\,\eta\geqslant f_1(\boldsymbol{x})\right\},\tag{3.3.11}$$

$$S_2=\left\{(\boldsymbol{x},\eta)\in\mathscr{H}\times R\,|\,-f_2(\boldsymbol{x})\geqslant\eta\right\}.\tag{3.3.12}$$

由 $(\operatorname{dom}f_1)\cap(\operatorname{dom}f_2)\neq\varnothing$ 和 f_1 是凸函数，根据定理 2.1.5 知 S_1 是凸集. 同理，由 f_2 是凸函数可推知 S_2 是非空凸集. 又由 f_1 在点 $\boldsymbol{x}^0$ 处是连续的，根据定理 2.2.1，它在 $\operatorname{int}(\operatorname{dom}f_1)$ 内是连续的，于是由推论 2.2.2 知 $\operatorname{int}S_1=\operatorname{int}(\operatorname{epi}f_1)\neq\varnothing$. 此外，可以证明有 $\operatorname{int}S_1\cap S_2=\varnothing$. 事实上，否则，若存在 $(\bar{\boldsymbol{x}},\bar{\eta})\in\operatorname{int}S_1\cap S_2$，则由(3.3.11)和(3.3.12)有

$$f_1(\bar{\boldsymbol{x}})<\bar{\eta}\leqslant-f_2(\bar{\boldsymbol{x}}).$$

根据此，导致

$$\inf_{\boldsymbol{y}\in\mathscr{H}}\left[f_1(\boldsymbol{y})+f_2(\boldsymbol{y})\right]\leqslant f_1(\bar{\boldsymbol{x}})+f_2(\bar{\boldsymbol{x}})<0,$$

它与(3.3.10)相矛盾. 根据以上得到的 $\operatorname{int}S_1\neq\varnothing$，$S_2\neq\varnothing$ 和 $\operatorname{int}S_1\cap S_2=\varnothing$，利用定理 1.3.7 的(2)，可知存在闭超平面严格分离 $\operatorname{int}S_1$ 和 S_2. 按定义 1.3.2 的(2)，也即存在 $(\boldsymbol{x}^*,\eta^*)\in(\mathscr{H}^*\times R)\setminus\{\boldsymbol{0}\}$ 有

$$\left\langle(\boldsymbol{x}^*,\eta^*),(\boldsymbol{x}^1,\eta_1)\right\rangle<\left\langle(\boldsymbol{x}^*,\eta^*),(\boldsymbol{x}^2,\eta_2)\right\rangle$$

$$\forall(\boldsymbol{x}^1,\eta_1)\in\operatorname{int}S_1,\ \forall(\boldsymbol{x}^2,\eta_2)\in S_2,$$

或

$$\langle \boldsymbol{x}^*, \boldsymbol{x}^1 \rangle + \eta^* \eta_1 < \langle \boldsymbol{x}^*, \boldsymbol{x}^2 \rangle + \eta^* \eta_2$$

$$\forall (\boldsymbol{x}^1, \eta_1) \in \mathrm{int} S_1, \quad \forall (\boldsymbol{x}^2, \eta_2) \in S_2. \qquad (3.3.13)$$

(3.3.13)左端的 η_1 可以任意大，故 $\eta^* \leqslant 0$. 但若有 $\eta^* = 0$，则从 $\boldsymbol{x}^0 \in (\mathrm{dom} f_1) \cap (\mathrm{dom} f_2)$ 且 f_1 在点 $\boldsymbol{x}^0$ 处是连续的可知，存在 $\varepsilon > 0$ 有

$$(\boldsymbol{x}^0, f_1(\boldsymbol{x}^0) + \varepsilon) \in \mathrm{int} S_1, (\boldsymbol{x}^0, -f_2(\boldsymbol{x}^0)) \in S_2,$$

于是由(3.3.13)导致矛盾：

$$\langle \boldsymbol{x}^*, \boldsymbol{x}^0 \rangle < \langle \boldsymbol{x}^*, \boldsymbol{x}^0 \rangle.$$

因此，我们得知 $\eta^* < 0$.

令 $\boldsymbol{x}_1^* = -\dfrac{\boldsymbol{x}^*}{\eta^*}$，由(3.3.13)注意到 $\eta^* < 0$，得

$$-\langle \boldsymbol{x}_1^*, \boldsymbol{x}^1 \rangle + \eta_1 \geqslant -\langle \boldsymbol{x}_1^*, \boldsymbol{x}^2 \rangle + \eta_2$$

$$\forall (\boldsymbol{x}^1, \eta_1) \in S_1, \forall (\boldsymbol{x}^2, \eta_2) \in S_2.$$

特别是，取 $\eta_1 = f_1(\boldsymbol{x}^1)$，$\eta_2 = -f_2(\boldsymbol{x}^2)$，则有

$$-\langle \boldsymbol{x}_1^*, \boldsymbol{x}^1 \rangle + f_1(\boldsymbol{x}^1) \geqslant -\langle \boldsymbol{x}_1^*, \boldsymbol{x}^2 \rangle - f_2(\boldsymbol{x}^2)$$

$$\forall \boldsymbol{x}^1 \in \mathrm{dom} f_1, \forall \boldsymbol{x}^2 \in \mathrm{dom} f_2. \qquad (3.3.14)$$

在(3.3.14)中取 $\boldsymbol{x}^2 = \boldsymbol{0}$，因为 $f_1(\boldsymbol{0}) = f_2(\boldsymbol{0}) = 0$，故得

$$\langle \boldsymbol{x}_1^*, \boldsymbol{x}^1 \rangle \leqslant f_1(\boldsymbol{x}^1) = f_1(\boldsymbol{x}^1) - f_1(\boldsymbol{0}) \quad \forall \boldsymbol{x}^1 \in \mathrm{dom} f_1.$$

据此，由定义 3.2.1 即知 $\boldsymbol{x}_1^* \in \partial f_1(\boldsymbol{0})$. 在(3.3.14)中再取 $\boldsymbol{x}^1 = \boldsymbol{0}$，同理有

$$-\langle \boldsymbol{x}_1^*, \boldsymbol{x}^2 \rangle \leqslant f_2(\boldsymbol{x}^2) = f_2(\boldsymbol{x}^2) - f_2(\boldsymbol{0}) \quad \forall \boldsymbol{x}^2 \in \mathrm{dom} f_2,$$

于是得 $-\boldsymbol{x}_1^* \in \partial f_2(\boldsymbol{0})$. 因此得到 $\boldsymbol{0} = \boldsymbol{x}_1^* - \boldsymbol{x}_1^* \in \partial f_1(\boldsymbol{0}) + \partial f_2(\boldsymbol{0})$. □

以下讨论次微分映射的连续性. 为此，先给出集值映射连续性

的定义.

设 $\mathscr{Y}$ 是线性拓扑空间.

定义 3.3.1 设点 $\boldsymbol{x}\in\mathscr{X}$，集合 $\psi(\boldsymbol{x})\subset\mathscr{Y}$，$\psi:\mathscr{X}\to 2^{\mathscr{Y}}$，$\boldsymbol{x}\mapsto\psi(\boldsymbol{x})$ 是集值映射.

(1) 若点 $\boldsymbol{x}^0\in\mathscr{X}$，并且对任意的点列 $\{\boldsymbol{x}^k\}\subset\mathscr{X}$，$\boldsymbol{x}^k\to\boldsymbol{x}^0$ $(k\to\infty)$，以及 $\boldsymbol{y}^k\in\psi(\boldsymbol{x}^k)$，$\boldsymbol{y}^k\to\boldsymbol{y}^0$，有 $\boldsymbol{y}^0\in\psi(\boldsymbol{x}^0)$，则称集值映射 ψ 在点 $\boldsymbol{x}^0\in\mathscr{X}$ 处是上半连续的. 若 ψ 在集合 $S\subset\mathscr{X}$ 的每一点处都是上半连续的，则称 ψ 是 S 上的上半连续集值映射，或 ψ 在 S 上是上半连续的.

(2) 若点 $\boldsymbol{x}^0\in\mathscr{X}$，并且对任意的点列 $\{\boldsymbol{x}^k\}\subset\mathscr{X}$，$\boldsymbol{x}^k\to\boldsymbol{x}^0$ $(k\to\infty)$，以及 $\boldsymbol{y}^0\in\psi(\boldsymbol{x}^0)$，存在正整数 M 和点列 $\{\boldsymbol{y}^k\}\subset\mathscr{Y}$ 有 $\boldsymbol{y}^k\in\psi(\boldsymbol{x}^k)$ $(k\geqslant M)$ 和 $\boldsymbol{y}^k\to\boldsymbol{y}^0$，则称集值映射 ψ 在点 $\boldsymbol{x}^0\in\mathscr{X}$ 处是下半连续的. 若 ψ 在集合 $S\subset\mathscr{X}$ 的每一点处都是下半连续的，则称 ψ 是 S 上的下半连续集值映射，或 ψ 在 S 上是下半连续的.

(3) 若 ψ 在点 $\boldsymbol{x}^0\in\mathscr{X}$ 处既是上半连续的又是下半连续的，则称集值映射 ψ 在点 $\boldsymbol{x}^0$ 处是连续的. 若 ψ 在集合 $S\subset\mathscr{X}$ 的每一点处都是连续的，则称 ψ 是 S 上的连续集值映射，或 ψ 在 S 上是连续的.

注 3.3.1 当 $\mathscr{Y}=R$，并且 $\psi=f:\mathscr{X}\to R$ 是实值函数时，定义 3.3.1 的(1)意即对任意的 $\{\boldsymbol{x}^k\}\subset\mathscr{X}$，$\boldsymbol{x}^k\to\boldsymbol{x}^0(k\to\infty)$ 有 $f(\boldsymbol{x}^k)\to f(\boldsymbol{x}^0)$，也即 $\lim\limits_{\boldsymbol{x}^k\to\boldsymbol{x}^0}f(\boldsymbol{x}^k)=f(\boldsymbol{x}^0)$. 由 $\{\boldsymbol{x}^k\}$ 的任意性，可知它等价于对任意的 $\varepsilon>0$ 存在 $\boldsymbol{x}^0$ 的邻域 $U(\boldsymbol{x}^0)\subset\mathscr{X}$，有

$$\left|f(\boldsymbol{x})-f(\boldsymbol{x}^0)\right|<\varepsilon\quad\forall\boldsymbol{x}\in U(\boldsymbol{x}^0).$$

对照定义 2.2.1，即 f 在点 $\boldsymbol{x}^0$ 处是连续的. 类似地，定义 3.3.1 的(2)意即对任意的 $\{\boldsymbol{x}^k\}\subset\mathscr{X}$，$\boldsymbol{x}^k\to\boldsymbol{x}^0(k\to\infty)$，存在 M 和数列 $\left\{f(\boldsymbol{x}^k)\right\}\subset R$，当 $k\geqslant M$ 时，有 $f(\boldsymbol{x}^k)\to f(\boldsymbol{x}^0)$，即 $\lim\limits_{\boldsymbol{x}^k\to\boldsymbol{x}^0}f(\boldsymbol{x}^k)=f(\boldsymbol{x}^0)$. 同理，它也即 f 在点 $\boldsymbol{x}^0$ 处是连续的. 因此，当集值映射退化为实值函数时，它的上半连续性和下半连续性(这时等同)以及连

续性都是实值函数(实泛函)的连续性. 但应指出,对于线性拓扑空间上的一般集值映射而言,它的上半连续性和下半连续性是互不包含的.

定理 3.3.5　设 $f:\mathscr{X}\to R\cup\{+\infty\}$ 是正常凸函数. 若 f 在点 $\boldsymbol{x}^0\in\operatorname{dom}f$ 处是连续的,则 ∂f 在 $\boldsymbol{x}^0$ 处是上半连续的.

证明　从 f 在点 $\boldsymbol{x}^0$ 处是连续的,由定义 2.2.1 知 f 在 $\boldsymbol{x}^0$ 的某邻域 $U_1(\boldsymbol{x}^0)$ 内上有界. 因此,由定理 2.2.1,它在 $U_1(\boldsymbol{x}^0)$ 内是连续的. 对于 $\boldsymbol{x}\in U_1(\boldsymbol{x}^0)$, 由定义 3.1.1 的(1)有

$$f'_+(\boldsymbol{x};\boldsymbol{d})=\lim_{t\to0^+}\frac{f(\boldsymbol{x}+t\boldsymbol{d})-f(\boldsymbol{x})}{t}\quad\forall\boldsymbol{d}\in\mathscr{X},\tag{3.3.15}$$

故对任何 $\boldsymbol{d}\in\mathscr{X}$, 函数

$$F(\boldsymbol{x},t)=\begin{cases}\dfrac{f(\boldsymbol{x}+t\boldsymbol{d})-f(\boldsymbol{x})}{t}, & t>0;\\ f'_+(\boldsymbol{x};\boldsymbol{d}), & t=0,\end{cases}$$

在 $U_1(\boldsymbol{x}^0)\times[0,\infty)$ 内是连续的. 据此,按定义 2.2.1 可知,对任意的 $\varepsilon>0$ 存在 $\boldsymbol{x}^0$ 的邻域 $U(\boldsymbol{x}^0)\subset U_1(\boldsymbol{x}^0)$ 和 $\delta_1>0$, 使得

$$\left|\frac{f(\boldsymbol{x}+t\boldsymbol{d})-f(\boldsymbol{x})}{t}-f'_+(\boldsymbol{x}^0;\boldsymbol{d})\right|<\frac{\varepsilon}{2}$$

$$\forall\boldsymbol{x}\in U(\boldsymbol{x}^0),\ \forall t\in(0,\delta_1).\tag{3.3.16}$$

另外,从(3.3.15)得知对上述的 $\varepsilon>0$ 又存在 $\delta_2>0$, 有

$$\left|\frac{f(\boldsymbol{x}^0+t\boldsymbol{d})-f(\boldsymbol{x}^0)}{t}-f'_+(\boldsymbol{x}^0;\boldsymbol{d})\right|<\frac{\varepsilon}{2}\quad\forall t\in(0,\delta_2).\tag{3.3.17}$$

取 $\delta=\min\{\delta_1,\delta_2\}$. 由(3.3.16)和(3.3.17)可得对任意的 $\boldsymbol{x}\in U(\boldsymbol{x}^0)$ 和任意的 $t\in(0,\delta)$, 有

$$\frac{f(\boldsymbol{x}+t\boldsymbol{d})-f(\boldsymbol{x})}{t}<f'_+(\boldsymbol{x}^0;\boldsymbol{d})+\frac{\varepsilon}{2}$$

$$< \frac{f(\boldsymbol{x}^0 + t\boldsymbol{d}) - f(\boldsymbol{x}^0)}{t} + \frac{\varepsilon}{2} + \frac{\varepsilon}{2},$$

即

$$\frac{f(\boldsymbol{x} + t\boldsymbol{d}) - f(\boldsymbol{x})}{t} < \frac{f(\boldsymbol{x}^0 + t\boldsymbol{d}) - f(\boldsymbol{x}^0)}{t} + \varepsilon$$

$$\forall \boldsymbol{x} \in U(\boldsymbol{x}^0),\ \forall t \in (0, \delta).$$

由此，根据定义 3.1.1 的(1)，对任意的 $\boldsymbol{x} \in U(\boldsymbol{x}^0)$ 有

$$\begin{aligned} f'_+(\boldsymbol{x}; \boldsymbol{d}) &= \lim_{t \to 0^+} \frac{f(\boldsymbol{x} + t\boldsymbol{d}) - f(\boldsymbol{x})}{t} \\ &\leqslant \lim_{t \to 0^+} \frac{f(\boldsymbol{x}^0 + t\boldsymbol{d}) - f(\boldsymbol{x}^0)}{t} + \varepsilon \\ &= f'_+(\boldsymbol{x}^0; \boldsymbol{d}) + \varepsilon, \end{aligned}$$

从而

$$\inf_{U(\boldsymbol{x}^0)} \sup_{\boldsymbol{x} \in U(\boldsymbol{x}^0)} f'_+(\boldsymbol{x}; \boldsymbol{d}) \leqslant f'_+(\boldsymbol{x}^0; \boldsymbol{d}) + \varepsilon.$$

令 $\varepsilon \to 0$，得到

$$\inf_{U(\boldsymbol{x}^0)} \sup_{\boldsymbol{x} \in U(\boldsymbol{x}^0)} f'_+(\boldsymbol{x}; \boldsymbol{d}) \leqslant f'_+(\boldsymbol{x}^0; \boldsymbol{d}). \qquad (3.3.18)$$

根据注 2.1.3，即知 $f'_+(\boldsymbol{x}; \boldsymbol{d})$ 关于 $\boldsymbol{x}$ 在 $\boldsymbol{x}^0$ 处是上半连续的.

因为 f 是正常凸函数并且在 $U(\boldsymbol{x}^0)$ 内是连续的，由定理 3.2.1 可知，对任意的 $\boldsymbol{x} \in U(\boldsymbol{x}^0)$ 有 $\partial f(\boldsymbol{x}) \neq \varnothing$. 设 $\boldsymbol{x}^* \in \partial f(\boldsymbol{x})$，根据定理 3.2.2 得到

$$\langle \boldsymbol{x}^*, \boldsymbol{d} \rangle \leqslant f'_+(\boldsymbol{x}; \boldsymbol{d}) \quad \forall \boldsymbol{d} \in \mathscr{X},\ \forall \boldsymbol{x} \in U(\boldsymbol{x}^0). \qquad (3.3.19)$$

以下证明集值映射 ∂f 在 $\boldsymbol{x}^0$ 处是上半连续的. 为此，任取点列 $\{\boldsymbol{x}_k\} \subset \mathscr{X}$，$\boldsymbol{x}_k \to \boldsymbol{x}^0$ $(k \to \infty)$，以及 $\boldsymbol{x}_k^* \in \partial f(\boldsymbol{x}_k)$，$\boldsymbol{x}_k^* \to \boldsymbol{x}^*$. 从 (3.3.19) 有

$$\langle \boldsymbol{x}_k^*, \boldsymbol{d} \rangle \leqslant f'_+(\boldsymbol{x}_k; \boldsymbol{d}) \quad \forall \boldsymbol{d} \in \mathscr{X},\ \forall \boldsymbol{x}_k \in U(\boldsymbol{x}^0).$$

据此，由(3.3.18)知，对任意的 $\boldsymbol{d} \in \mathscr{X}$ 有

$$\langle \boldsymbol{x}^*, \boldsymbol{d} \rangle = \lim_{k\to\infty} \langle \boldsymbol{x}_k^*, \boldsymbol{d} \rangle \leqslant \inf_{U(\boldsymbol{x}^0)} \sup_{\boldsymbol{x}_k \in U(\boldsymbol{x}^0)} f'_-(\boldsymbol{x}_k; \boldsymbol{d})$$

$$\leqslant \inf_{U(\boldsymbol{x}^0)} \sup_{\boldsymbol{x} \in U(\boldsymbol{x}^0)} f'_+(\boldsymbol{x}; \boldsymbol{d})$$

$$\leqslant f'_+(\boldsymbol{x}^0; \boldsymbol{d}),$$

于是根据定理 3.2.2 可知 $\boldsymbol{x}^* \in \partial f(\boldsymbol{x}^0)$. 最后，由定义 3.3.1 的(1)，便得 ∂f 在 $\boldsymbol{x}^0$ 处是上半连续的. □

最后，阐述次微分映射的单调性.

定义 3.3.2　设点 $\boldsymbol{x} \in \mathscr{X}$，集合 $\psi(\boldsymbol{x}) \subset \mathscr{X}^*$，$\psi: \mathscr{X} \to 2^{\mathscr{X}^*}$，$\boldsymbol{x} \mapsto \psi(\boldsymbol{x})$ 是集值映射. 若对任意的 $\boldsymbol{x}, \boldsymbol{y} \in \mathscr{X}$ 有

$$\langle \boldsymbol{y}^* - \boldsymbol{x}^*, \boldsymbol{y} - \boldsymbol{x} \rangle \geqslant 0 \quad \forall \boldsymbol{x}^* \in \psi(\boldsymbol{x}), \forall \boldsymbol{y}^* \in \psi(\boldsymbol{y}), \tag{3.3.20}$$

则称 ψ 是 $\mathscr{X}$ 上的单调集值映射，或称集值映射 ψ 在 $\mathscr{X}$ 上是单调的.

注 3.3.2　当 $\mathscr{X} = R$ 和 $2^{\mathscr{Y}} = R$，ψ 是单变量函数 $\psi = f: R \to R$ 时，定义 3.3.2 中的 $x^* = f(x)$，$y^* = f(y)$，(3.3.20)即为

$$\langle f(y) - f(x), y - x \rangle \geqslant 0.$$

由此，当 $x \leqslant y$ 时有 $f(x) \leqslant f(y)$，这也即 f 是 R 上的单调函数.

定理 3.3.6　设 $f: \mathscr{X} \to R \cup \{+\infty\}$ 是连续的正常凸函数，则 ∂f 在 $\mathscr{X}$ 上是单调的.

证明　对于任意的 $\boldsymbol{x}, \boldsymbol{y} \in \mathscr{X}$，任取 $\boldsymbol{x}^* \in \partial f(\boldsymbol{x})$ 和 $\boldsymbol{y}^* \in \partial f(\boldsymbol{y})$，由定义 3.2.1 有

$$\langle \boldsymbol{x}^*, \boldsymbol{y} - \boldsymbol{x} \rangle \leqslant f(\boldsymbol{y}) - f(\boldsymbol{x}) \quad \forall \boldsymbol{y} \in \mathscr{X} \tag{3.3.21}$$

和

$$\langle \boldsymbol{y}^*, \boldsymbol{x} - \boldsymbol{y} \rangle \leqslant f(\boldsymbol{x}) - f(\boldsymbol{y}) \quad \forall \boldsymbol{x} \in \mathscr{X}. \tag{3.3.22}$$

由(3.3.22)有 $\langle -\boldsymbol{y}^*, \boldsymbol{y} - \boldsymbol{x} \rangle \leqslant f(\boldsymbol{x}) - f(\boldsymbol{y})$. 将它与(3.3.21)的两端相加得$\langle \boldsymbol{x}^* - \boldsymbol{y}^*, \boldsymbol{y} - \boldsymbol{x} \rangle \leqslant 0$，注意，$\boldsymbol{x}^* \in \partial f(\boldsymbol{x})$ 和 $\boldsymbol{y}^* \in$

$\partial f(y)$ 是任取的，有

$$\langle x^* - y^*, y - x\rangle \geqslant 0 \quad \forall x^* \in \partial f(x), \forall y^* \in \partial f(y).$$

因此，由定义 3.3.2 即知 ∂f 在 $\mathscr{X}$ 上是单调的. ▯

进而，下面考虑更强的单调性.

定义 3.3.3 设点 $x \in \mathscr{X}$，集合 $\psi(x) \subset \mathscr{X}^*$，$\psi: \mathscr{X} \to 2^{\mathscr{X}^*}$，$x \mapsto \psi(x)$ 是集值映射. 若对任何正整数 $m \geqslant 2$ 和任意的 $x^i \in \mathscr{X}$ $(i = 1, \cdots, m)$ 有

$$\langle x_1^*, x^2 - x^1\rangle + \langle x_2^*, x^3 - x^2\rangle + \cdots + \langle x_m^*, x^1 - x^m\rangle$$

$$\leqslant 0 \quad \forall x_1^* \in \psi(x^1), \cdots, \forall x_m^* \in \psi(x^m), \qquad (3.3.23)$$

则称 ψ 是 $\mathscr{X}$ 上的循环单调集值映射，或集值映射 ψ 在 $\mathscr{X}$ 上是循环单调的.

显然，在定义 3.3.3 中当 $m = 2$ 时，(3.3.23)即

$$\langle x_1^*, x^2 - x^1\rangle + \langle x_2^*, x^1 - x^2\rangle \leqslant 0$$

$$\forall x_1^* \in \psi(x^1), \forall x_2^* \in \psi(x^2),$$

或

$$\langle x_2^* - x_1^*, x^2 - x^1\rangle \geqslant 0 \quad \forall x_1^* \in \psi(x^1), \forall x_2^* \in \psi(x^2).$$

于是，由定义 3.3.2 可知，这时 ψ 在 $\mathscr{X}$ 上是单调的.

定理 3.3.7 设 $f: \mathscr{X} \to R \cup \{+\infty\}$ 是连续的正常凸函数，则 ∂f 在 $\mathscr{X}$ 上是循环单调的.

证明 对任意的 $x^i \in \mathscr{X} (i = 1, \cdots, m)$，任取 $x_i^* \in \partial f(x^i)$，由定义 3.2.1 有

$$\langle x_1^*, x^2 - x^1\rangle \leqslant f(x^2) - f(x^1),$$

$$\langle x_2^*, x^3 - x^2\rangle \leqslant f(x^3) - f(x^2),$$

$$\cdots\cdots\cdots\cdots$$

$$\langle x_m^*, x^1 - x^m\rangle \leqslant f(x^1) - f(x^m).$$

将以上各式两端相加，注意各 $x_i^* \in \partial f(x^i)$ 是任取的，即得 (3.3.23). 由此，按定义 3.3.3 即可知 ∂f 在 $\mathscr{X}$ 上是循环单调的. □

§ 3.4　凸规划的最优性条件

利用凸函数的次微分，本节论述凸规划问题的最优解要满足的最优性条件. 这是数学规划理论中最基本的问题.

设 $\mathscr{X}$ 是线性空间，$f: \mathscr{X} \to R \cup \{\pm\infty\}$，$g_i: \mathscr{X} \to R \cup \{\pm\infty\}$ $(i = 1, \cdots, p)$ 和 $h_j: \mathscr{X} \to R \cup \{\pm\infty\}$ $(j = 1, \cdots, q)$ 都是广义实值函数，则一般形式的数学规划问题记作

$$\begin{cases} \min f(\boldsymbol{x}) \\ \text{s.t. } g_i(\boldsymbol{x}) \leqslant 0, \ i = 1, \cdots, p, \\ \qquad h_j(\boldsymbol{x}) = 0, \ j = 1, \cdots, q. \end{cases} \tag{MP}$$

上式中的 f 称为(MP)的**目标函数**，$\min f(x)$ 表示对 $f(\boldsymbol{x})$ 的极小化；$g_i(i = 1, \cdots, p)$ 和 $h_j(j = 1, \cdots, q)$ 称为**约束函数**，s.t. 表示受约束于，而称

$$g_i(\boldsymbol{x}) \leqslant 0, \ i = 1, \cdots, p$$

是(MP)的不等式约束条件；

$$h_j(\boldsymbol{x}) = 0, \ j = 1, \cdots, q$$

是(MP)的等式约束条件. 记

$$X = \{\boldsymbol{x} \in \mathscr{X} \mid g_i(\boldsymbol{x}) \leqslant 0, \ i = 1, \cdots, p; \ h_j(\boldsymbol{x}) = 0, \ j = 1, \cdots, q\}, \tag{3.4.1}$$

则问题(MP)也可简记作

$$\min_{x \in X} f(\boldsymbol{x}), \tag{MP}$$

其中 $X \subset \mathscr{V}$ 称为(MP)的约束集.

特别是,当 $X = \mathscr{V}$ 时,称(MP)是无约束极小化问题,并记它为

$$\min f(\boldsymbol{x}). \tag{UMP}$$

当 $X \neq \varnothing$ 是 $\mathscr{V}$ 的真子集时,则称(MP)是约束极小化问题.

若问题(MP)中的目标函数 f 是凸函数,并且约束集 X 是凸集,则称它是凸规划问题或凸规划. 若(MP)中的目标函数 f 以及约束函数 $g_i(i=1,\cdots,p)$ 和 $h_j(j=1,\cdots,q)$ 都是线性函数,则称它是线性规划问题或线性规划.

注 3.4.1 若(MP)中的 f, $g_i(i=1,\cdots,p)$ 是凸函数,h_j $(j=1,\cdots,q)$ 是仿射函数,则(MP)是一凸规划. 事实上,记

$$S_1 = \bigcap_{i=1}^{p}\{\boldsymbol{x}\in\mathscr{V}\mid g_i(\boldsymbol{x})\leqslant 0\},$$

$$S_2 = \bigcap_{j=1}^{q}\{\boldsymbol{x}\in\mathscr{V}\mid h_j(\boldsymbol{x}) = 0\},$$

由(3.4.1)知

$$X = S_1 \cap S_2.$$

因为 $g_i(i=1,\cdots,p)$ 和 $h_j(j=1,\cdots,q)$ 都是凸函数,根据定理 2.1.2 知,它们相应的水平集 $\{\boldsymbol{x}\in\mathscr{V}\mid g_i(\boldsymbol{x})\leqslant 0\}(i=1,\cdots,p)$ 和 $\{\boldsymbol{x}\in\mathscr{V}\mid h_j(\boldsymbol{x})=0\}(j=1,\cdots,q)$ 都是凸集,所以由定理 1.1.2的(2)得知它们的交集 X 是凸集. 此外,由于线性函数也是凸函数,故线性规划也是凸规划.

对于凸规划问题(MP),考虑它的约束集 X 上的指示函数

$$\delta_X(\boldsymbol{x}) = \begin{cases} 0, & \boldsymbol{x}\in X; \\ +\infty, & \boldsymbol{x}\in\mathscr{V}\setminus X, \end{cases}$$

则约束极小化问题(MP)可以转化为等价的无约束极小化问题:

$$\min\left[f(\boldsymbol{x}) + \delta_X(\boldsymbol{x})\right]. \tag{3.4.2}$$

定义 3.4.1　设集合 $X \subset \mathscr{Y}$ 非空，$f: \mathscr{Y} \to R \cup \{\pm\infty\}$ 是广义实值函数. 若 $\tilde{x} \in X$，并且

$$f(\tilde{x}) \leqslant f(x) \quad \forall x \in X,$$

则称 $\tilde{x}$ 是数学规划问题(MP)的最优解(或极小解)，$f(\tilde{x})$ 是(MP)的最优值(或极小值). (MP)的所有最优解组成的集合称为(MP)的最优解集，记作 $\tilde{X}$.

对于线性拓扑空间 $\mathscr{X}$ 上的无约束凸规划问题(UMP)，它的最优解有如下简单的最优性充分必要条件.

定理 3.4.1　设 $f: \mathscr{X} \to R \cup \{+\infty\}$ 是正常凸函数，则 $\tilde{x}$ 是(UMP)的最优解当且仅当 $\mathbf{0} \in \partial f(\tilde{x})$.

证明　由定义 3.4.1，注意到这时 $X = \mathscr{X}$，则 $\tilde{x}$ 是(UMP)的最优解意味着

$$f(x) - f(\tilde{x}) \geqslant 0 \quad \forall x \in \mathscr{X},$$

即

$$\langle \mathbf{0}, x - \tilde{x} \rangle \leqslant f(x) - f(\tilde{x}) \quad \forall x \in \mathscr{X}.$$

依据定义 3.2.1，它与 $\mathbf{0} \in \partial f(\tilde{x})$ 等价. □

推论 3.4.2　设 $f: R^n \to R$ 是正常凸函数. 若 f 在 R^n 上可微，则 $\tilde{x}$ 是(UMP)的最优解当且仅当 $\nabla f(\tilde{x}) = \mathbf{0}$.

证明　因为 f 是 R^n 上的可微函数，由推论 3.2.5 可知，对任意的 $x \in R^n$，有 $\partial f(x) = \{\nabla f(x)\}$. 由此，根据定理 3.4.1 即可得到 $\nabla f(\tilde{x}) = \mathbf{0}$. □

下面的定理指出，凸规划问题的最优解集必是凸集.

定理 3.4.3　设 $X \subset \mathscr{X}$ 是非空凸集，$f: \mathscr{X} \to R \cup \{+\infty\}$ 是正常凸函数，则(MP)的最优解集 $\tilde{X}$ 是凸集.

证明　设 $\tilde{X} \neq \varnothing$，则有最优解 $\tilde{x} \in \tilde{X}$. 记(MP)的最优值 $\tilde{f} = f(\tilde{x})$，由定义 3.4.1 知

$$\tilde{X} = \{x \in X \mid f(\tilde{x}) \leqslant \tilde{f}\}.$$

因为 X 是非空凸集，f 是凸函数，所以由定理 2.1.2 得知，对于 $\tilde{f} \in R$，其相应的水平集 $\tilde{X}$ 是凸集. □

为论述约束凸规划问题(MP)的最优解要满足的最优性条件，我们先给出如下的引理.

引理 3.4.4 设 $F: \mathscr{X} \to R \cup \{+\infty\}$ 是正常凸函数，$h_j: \mathscr{X} \to R(j=1, \cdots, q)$ 是仿射函数，$\boldsymbol{h}(\boldsymbol{x}) = (h_1(\boldsymbol{x}), \cdots, h_q(\boldsymbol{x}))^T$，并且 $\boldsymbol{0} \in (\boldsymbol{h}(\mathrm{dom}F))^i$，则

$$F(\boldsymbol{x}) \geqslant 0 \quad \forall \boldsymbol{x} \in \{\boldsymbol{x} \in \mathscr{X} \mid \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}\} \tag{3.4.3}$$

当且仅当存在 $w > 0$ 和 $\mu_j \in R(j=1, \cdots, q)$ 使得

$$wF(\boldsymbol{x}) + \sum_{j=1}^{q} \mu_j h_j(\boldsymbol{x}) \geqslant 0 \quad \forall \boldsymbol{x} \in \mathrm{dom}F. \tag{3.4.4}$$

证明 充分性. 由 $\boldsymbol{0} \in (\boldsymbol{h}(\mathrm{dom}F))^i$ 可知存在 $\boldsymbol{x} \in \mathrm{dom}F$ 使 $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$，据此，从(3.4.4)我们有

$$wF(\boldsymbol{x}) \geqslant 0 \quad \forall \boldsymbol{x} \in \{\boldsymbol{x} \in \mathscr{X} \mid \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}\}.$$

因为 $w > 0$，所以得到(3.4.3).

必要性. 作向量函数 $\boldsymbol{G}: \mathrm{dom}F \to R \times R^q$，

$$\boldsymbol{G}(\boldsymbol{x}) = (F(\boldsymbol{x}), \boldsymbol{h}(\boldsymbol{x})), \tag{3.4.5}$$

记

$$A = \{(y, \boldsymbol{z}) \in R \times R^q \mid y \geqslant 0, \boldsymbol{z} = \boldsymbol{0}\}. \tag{3.4.6}$$

由(3.4.5)和(3.4.3)易知集合 $\boldsymbol{G}(\mathrm{dom}F) + A \subset R^{q+1}$ 非空. 我们指出，$\boldsymbol{G}(\mathrm{dom}F) + A$ 还是凸的. 事实上，任取 $(y^1, \boldsymbol{z}^1), (y^2, \boldsymbol{z}^2) \in \boldsymbol{G}(\mathrm{dom}F) + A$，由(3.4.5)和(3.4.6)可知存在 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in \mathrm{dom}F$，使

$$F(\boldsymbol{x}^1) \leqslant y^1, \ F(\boldsymbol{x}^2) \leqslant y^2,$$

$$\boldsymbol{h}(\boldsymbol{x}^1) = \boldsymbol{z}^1, \ \boldsymbol{h}(\boldsymbol{x}^1) = \boldsymbol{z}^2.$$

于是，对任意的 $\lambda \in (0, 1)$，由于 f 是凸函数和 $h_j (j = 1, \cdots, q)$ 是仿射函数，我们有

$$F\left(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2\right) \leqslant \lambda F(\boldsymbol{x}^1) + (1-\lambda)F(\boldsymbol{x}^2)$$

$$\leqslant \lambda y^1 + (1-\lambda)y^2,$$

$$\boldsymbol{h}\left(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2\right) \leqslant \lambda \boldsymbol{h}(\boldsymbol{x}^1) + (1-\lambda)\boldsymbol{h}(\boldsymbol{x}^2)$$

$$\leqslant \lambda \boldsymbol{z}^1 + (1-\lambda)\boldsymbol{z}^2.$$

据此，得到 $(\lambda y^1 + (1-\lambda)y^2, \lambda \boldsymbol{z}^1 + (1-\lambda)\boldsymbol{z}^2) \in \boldsymbol{G}(\mathrm{dom}F) + A$，故由定义 1.1.1 知 $\boldsymbol{G}(\mathrm{dom}F) + A$ 是凸集.

另一方面，对于非空凸集

$$D = \left\{(y, \boldsymbol{z}) \in R \times R^q \,|\, y < 0, \boldsymbol{z} = \boldsymbol{0}\right\}, \qquad (3.4.7)$$

由(3.4.3)和(3.4.6)可知

$$D \not\subset \boldsymbol{G}(\mathrm{dom}F) + A.$$

由于任何有限维非空凸集的相对内部是非空的，故有

$$(\boldsymbol{G}(\mathrm{dom}F) + A)^{ri} \neq \varnothing, \; D^{ri} \neq \varnothing.$$

因此，凸集 $\boldsymbol{G}(\mathrm{dom}F) + A$ 和 D 是可分离的. 根据定理 1.3.5 的(1)可知，存在 $(w, \boldsymbol{\mu}) \in R \times R^q \backslash \{\boldsymbol{0}\}$，使得

$$\langle w, y' \rangle + \langle \boldsymbol{\mu}, \boldsymbol{z}' \rangle \leqslant \langle w, y'' \rangle + \langle \boldsymbol{\mu}, \boldsymbol{z}'' \rangle$$

$$\forall (y', \boldsymbol{z}') \in D \quad \forall (y'', \boldsymbol{z}'') \in \boldsymbol{G}(\mathrm{dom}F) + A.$$

记 $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_q)^T \in R^q$，在上式中取 $y'' = F(\boldsymbol{x})$，$\boldsymbol{z}'' = \boldsymbol{h}(\boldsymbol{x})$ $(\boldsymbol{x} \in \mathrm{dom}F)$，注意 $y' < 0$ 和 $\boldsymbol{z}' = \boldsymbol{0}$（见(3.4.7)），则有

$$\langle w, y' \rangle \leqslant \langle w, F(\boldsymbol{x}) \rangle + \langle \boldsymbol{\mu}, \boldsymbol{h}(\boldsymbol{x}) \rangle$$

$$\forall y' \in (-\infty, 0), \forall \boldsymbol{x} \in \mathrm{dom}F, \qquad (3.4.8)$$

即

$$wy' \leqslant wF(\boldsymbol{x}) + \sum_{j=1}^{q}\mu_j h_j(\boldsymbol{x})$$

$$\forall y' \in (-\infty, 0), \forall \boldsymbol{x} \in \mathrm{dom}F.$$

令 $y' \to 0$, 得到

$$wF(\boldsymbol{x}) + \sum_{j=1}^{q}\mu_j h_j(\boldsymbol{x}) \geqslant 0 \quad \forall \boldsymbol{x} \in \mathrm{dom}F,$$

这就是(3.4.4). 下面说明 $w > 0$. 首先,因为(3.4.8)中的 y' 可趋于 $-\infty$ 而其右端是有限数,故 w 不能取负值. 此外,由于 $\boldsymbol{0} \in (\boldsymbol{h}(\mathrm{dom}F))^i$,则知 $w \neq 0$. 因为否则,若 $w = 0$, 则由(3.4.8),将有

$$\langle \mu, \boldsymbol{h}(\boldsymbol{x}) \rangle \geqslant 0 \quad \forall \boldsymbol{x} \in \mathrm{dom}F,$$

即

$$\langle \mu, \boldsymbol{z} \rangle \geqslant 0 \quad \forall \boldsymbol{z} \in \boldsymbol{h}(\mathrm{dom}F).$$

因为 $\boldsymbol{0} \in (\boldsymbol{h}(\mathrm{dom}F))^i$,故上式仅当 $\mu = \boldsymbol{0}$ 时才有可能,这导致与 $(w, \mu) \neq \boldsymbol{0}$ 相矛盾,从而引理得证. ▯

下面阐述约束凸规划问题(MP)的最优解要满足的最优性条件.

定理 3.4.5 设 $f: \mathscr{X} \to R \cup \{+\infty\}$ 和 $g_i: \mathscr{X} \to R \cup \{+\infty\}(i = 1, \cdots, p)$ 是正常凸函数, $h_j: \mathscr{X} \to R(j = 1, \cdots, q)$ 是连续的仿射函数, $X \neq \varnothing$ 由(3.4.1)确定(其中 $\mathscr{Y} = \mathscr{X}$). 若 f 和 $g_i(i = 1, \cdots, p)$ 在点 $\bar{\boldsymbol{x}} \in X$ 处有限并且是连续的,则 $\bar{\boldsymbol{x}} \in \bar{X}$ 当且仅当存在 $\bar{w} \geqslant 0$, $(\bar{\lambda}_1, \cdots, \bar{\lambda}_p)^T \in R_+^p$ 和 $(\bar{\mu}_1, \cdots, \bar{\mu}_q)^T \in R^q$ 使得

$$\begin{cases} \boldsymbol{0} \in \bar{w}\,\partial f(\bar{\boldsymbol{x}}) + \sum\limits_{i=1}^{p} \bar{\lambda}_i \partial g_i(\bar{\boldsymbol{x}}) + \sum\limits_{j=1}^{q} \bar{\mu}_j \partial h_j(\bar{\boldsymbol{x}}); \\ \bar{\lambda}_i g_i(\bar{\boldsymbol{x}}) = 0, \ i = 1, \cdots, p; \\ (\bar{w}, \bar{\lambda}_1, \cdots, \bar{\lambda}_p)^T \neq \boldsymbol{0}. \end{cases} \tag{3.4.9}$$

证明 记 $\boldsymbol{h} = (h_1, \cdots, h_q)^T$,先设 $\boldsymbol{h}: \mathscr{X} \to R^q$ 是满射的,即 $\boldsymbol{h}(\mathscr{X}) = R^q$[19]. 因为 f 和 $g_i(i = 1, \cdots, p)$ 在点 $\bar{\boldsymbol{x}}$ 处有限且是连续

的，故有

$$\bar{x} \in (\mathrm{dom} f)^i \cap \left(\bigcap_{i=1}^{p} (\mathrm{dom} g_i)^i \right).$$

从 $\bar{x} \in X$，由(3.4.1)知 $\boldsymbol{h}(\bar{x}) = \boldsymbol{0}$，因为 $\boldsymbol{h}$ 是仿射向量函数和满射的，得知

$$\boldsymbol{0} = \boldsymbol{h}(\bar{x}) \in \left[\boldsymbol{h}(\mathrm{dom} f) \cap \left(\bigcap_{i=1}^{p} (\mathrm{dom} g_i) \right) \right]^i. \quad (3.4.10)$$

由 $\bar{x} \in \bar{X}$，根据定义 3.4.1 有

$$f(\boldsymbol{x}) - f(\bar{\boldsymbol{x}}) \geqslant 0 \quad \forall \boldsymbol{x} \in \mathscr{X},$$

按(3.4.1)（当 $\mathscr{Y} = \mathscr{X}$）得

$$\max_{1 \leqslant i \leqslant p} \left[f(\boldsymbol{x}) - f(\bar{\boldsymbol{x}}),\ g_i(\boldsymbol{x}) \right] \geqslant 0$$

$$\forall \boldsymbol{x} \in \{ \boldsymbol{x} \in \mathscr{X} \mid \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \}. \quad (3.4.11)$$

令

$$F(\boldsymbol{x}) = \max_{1 \leqslant i \leqslant p} \left[f(\boldsymbol{x}) - f(\bar{\boldsymbol{x}}),\ g_i(\boldsymbol{x}) \right], \quad (3.4.12)$$

由(3.4.11)我们有

$$F(\boldsymbol{x}) \geqslant 0 \quad \forall \boldsymbol{x} \in \{ \boldsymbol{x} \in \mathscr{X} \mid \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \}, \quad (3.4.13)$$

并且

$$\mathrm{dom} F = \mathrm{dom} f \cap \left(\bigcap_{i=1}^{p} (\mathrm{dom} g_i) \right).$$

此外，从(3.4.10)知

$$\boldsymbol{0} \in \left(\boldsymbol{h}(\mathrm{dom} F) \right)^i,$$

于是由引理 3.4.4 得知(3.4.13)成立的充分必要条件是，存在 $w > 0$ 和 $\bar{\mu}_j (j = 1, \cdots, q)$，使得

$$wF(\boldsymbol{x}) + \sum_{j=1}^{q} \bar{\mu}_j h_j(\boldsymbol{x}) \geqslant 0 \quad \forall \boldsymbol{x} \in \mathrm{dom} F.$$

由于 $w > 0$，上式等价于

$$wF(\boldsymbol{x}) + \sum_{j=1}^{q} \tilde{\mu}_j h_j(\boldsymbol{x}) \geqslant 0 \quad \forall \boldsymbol{x} \in \mathscr{X}. \qquad (3.4.14)$$

因为 $E(\boldsymbol{x}) = wF(\boldsymbol{x}) + \sum_{j=1}^{q} \tilde{\mu}_j h_j(\boldsymbol{x})$ 也是 $\mathscr{X}$ 上的正常凸函数，由(3.4.13)可知(3.4.14)意味着它在 $\tilde{\boldsymbol{x}} \in \mathscr{X}$ 处达到最优值 $E(\tilde{\boldsymbol{x}}) = 0$，所以根据定理 3.4.1，有

$$\boldsymbol{0} \in \partial E(\tilde{\boldsymbol{x}}) = \partial \left[wF + \sum_{j=1}^{q} \tilde{\mu}_j h_j \right](\tilde{\boldsymbol{x}}).$$

据此，由定理 3.3.4，从上式得到

$$\boldsymbol{0} \in w\partial F(\tilde{\boldsymbol{x}}) + \sum_{j=1}^{q} \tilde{\mu}_j \partial h_j(\tilde{\boldsymbol{x}}). \qquad (3.4.15)$$

另外，由(3.4.12)根据定理 3.3.3 有

$$\partial F(\tilde{\boldsymbol{x}}) = \mathrm{co}\left\{ \partial f(\tilde{\boldsymbol{x}}), \bigcup_{i \in I(\tilde{\boldsymbol{x}})} \partial g_i(\tilde{\boldsymbol{x}}) \right\},$$

其中

$$I(\tilde{\boldsymbol{x}}) = \{ i \in \{1, \cdots, p\} \mid g_i(\tilde{\boldsymbol{x}}) = 0 \}. \qquad (3.4.16)$$

因此，由定理 1.1.8 知，存在

$$\lambda_0 \geqslant 0,\ \lambda_i \geqslant 0\ \left(i \in I(\tilde{\boldsymbol{x}})\right),\ \lambda_i = 0\ \left(i \notin I(\tilde{\boldsymbol{x}})\right),\ \sum_{i=0}^{p} \lambda_i = 1, \qquad (3.4.17)$$

使得

$$\partial F(\tilde{\boldsymbol{x}}) = \lambda_0 \partial f(\tilde{\boldsymbol{x}}) + \sum_{i=1}^{p} \lambda_i \partial g_i(\tilde{\boldsymbol{x}}).$$

将它代入(3.4.15)有

$$\boldsymbol{0} \in w\lambda_0 \partial f(\tilde{\boldsymbol{x}}) + w \sum_{i=1}^{p} \lambda_i \partial g_i(\tilde{\boldsymbol{x}}) + \sum_{j=1}^{q} \tilde{\mu}_j \partial h_j(\tilde{\boldsymbol{x}}).$$

令 $\tilde{w} = w\lambda_0$，$\tilde{\lambda}_i = w\lambda_i (i = 1, \cdots, p)$，从上式即得(3.4.9)的第 1 式. 再由(3.4.16)和(3.4.17)可得(3.4.9)的第 2 式和第 3 式.

以上过程是可逆的，故所得结果是 $\bar{x} \in \tilde{X}$ 的充分必要条件.

若 $\boldsymbol{h}: \mathscr{X} \to R^q$ 不是满射的，则可取 $\boldsymbol{h}(\mathscr{X}) \neq \varnothing$ 来替代 R^q，作同样的讨论，即可使本定理得证. ▯

定理 3.4.5 称为凸规划的 Fritz John 定理，其中的(3.4.9)称为凸规划问题(MP)的最优解的 Fritz John 条件.

推论 3.4.6　设 $f: R^n \to R \cup \{+\infty\}$ 和 $g_i: R^n \to R \cup \{+\infty\} (i = 1, \cdots, p)$ 是正常凸函数，$h_j: R^n \to R (j = 1, \cdots, q)$ 是线性函数，$X \neq \varnothing$ 由(3.4.1)确定(其中 $\mathscr{X} = R^n$). 若 f 和 $g_i (i = 1, \cdots, p)$ 在点 $\bar{x} \in X$ 处可微，则 $\bar{x} \in \tilde{X}$ 当且仅当存在 $\tilde{w} \geqslant 0$，$(\tilde{\lambda}_1, \cdots, \tilde{\lambda}_p)^T \in R^p_+$ 和 $(\tilde{\mu}_1, \cdots, \tilde{\mu}_q)^T \in R^q$，使得

$$\begin{cases} \tilde{w}\, \nabla f(\bar{x}) + \sum_{i=1}^{p} \tilde{\lambda}_i\, \nabla g_i(\bar{x}) + \sum_{j=1}^{q} \tilde{\mu}_j\, \nabla h_j(\bar{x}) = 0; \\ \tilde{\lambda}_i g_i(\bar{x}) = 0,\ i = 1, \cdots, p; \\ (\tilde{w}, \tilde{\lambda}_1, \cdots, \tilde{\lambda}_p)^T \neq \boldsymbol{0}. \end{cases}$$

证明　由推论 3.2.5，这时 $\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$，$\partial g_i(\bar{x}) = \{\nabla g_i(\bar{x})\} (i = 1, \cdots, p)$ 和 $\partial h_j(\bar{x}) = \{\nabla h_j(\bar{x})\} (j = 1, \cdots, q)$. 因此，根据定理 3.4.5 即可得证. ▯

最后，介绍著名的 Kuhn-Tucker 条件.

定义 3.4.2　设 $g_i: \mathscr{X} \to R \cup \{\pm\infty\} (i = 1, \cdots, p)$ 和 $h_j: \mathscr{X} \to R \cup \{\pm\infty\} (j = 1, \cdots, q)$ 是广义实值函数，X 由(3.4.1)确定. 若存在点 $\bar{x} \in X$，使得

$$g_i(\bar{x}) < 0,\ i = 1, \cdots, p;\ h_j(\bar{x}) = 0,\ j = 1, \cdots, q, \tag{3.4.18}$$

则称(MP)的约束集 X(或约束函数 $g_i (i = 1, \cdots, p)$，$h_j (j = 1, \cdots, q)$)满足 Slater 约束规格.

定理 3.4.7 设 $f: \mathscr{X} \to R \cup \{+\infty\}$ 和 $g_i: \mathscr{X} \to R \cup \{+\infty\}(i=1, \cdots, p)$ 是正常凸函数，$h_j: \mathscr{X} \to R(j=1, \cdots, q)$ 是连续的仿射函数，$X \neq \varnothing$ 由(3.4.1)确定(其中 $\mathscr{Y} = \mathscr{X}$)．若 f 和 $g_i(i=1, \cdots, p)$ 在点 $\bar{x} \in X$ 处有限并且是连续的，X 满足 Slater 约束规格，则 $\bar{x} \in \widetilde{X}$ 当且仅当存在 $(\bar{\lambda}_1, \cdots, \bar{\lambda}_p)^T \in R^p_+$ 和 $(\bar{\mu}_1, \cdots, \bar{\mu}_q)^T \in R^q$，使得

$$\begin{cases} \boldsymbol{0} \in \partial f(\bar{x}) + \sum_{i=1}^{p} \bar{\lambda}_i \partial g_i(\bar{x}) + \sum_{j=1}^{q} \bar{\mu}_j \partial h_j(\bar{x}); \\ \bar{\lambda}_i g_i(\bar{x}) = 0, \ i = 1, \cdots, p. \end{cases} \tag{3.4.19}$$

证明 充分性．由定理 3.4.5，取 $\bar{w} = 1$ 即得．

必要性．根据定理 3.4.5，只要在条件(3.4.18)下证明其中的 $\bar{w} > 0$ 即可．事实上，若假设 $\bar{w} = 0$，则由(3.4.9)的第 1 式，有

$$\boldsymbol{0} \in \sum_{i=1}^{p} \bar{\lambda}_i \partial g_i(\bar{x}) + \sum_{j=1}^{q} \bar{\mu}_j \partial h_j(\bar{x}),$$

由定理 3.3.4 即

$$\boldsymbol{0} \in \partial \left[\sum_{i=1}^{p} \bar{\lambda}_i g_i + \sum_{j=1}^{q} \bar{\mu}_j h_j \right](\bar{x}).$$

据此，利用定理 3.4.1，得知 $\bar{x} \in X$ 是以 $\sum_{i=1}^{p} \bar{\lambda}_i g_i(x) + \sum_{j=1}^{q} \bar{\mu}_j h_j(x)$ 为目标函数的无约束极小化问题(P)的最优解．另外，由(3.4.18)、(3.4.16)和(3.4.17)得知，存在点 $\bar{x} \in X$ 使得

$$\sum_{i=1}^{p} \bar{\lambda}_i g_i(\bar{x}) + \sum_{j=1}^{q} \bar{\mu}_j h_j(\bar{x}) < 0 = \sum_{i=1}^{p} \bar{\lambda}_i g_i(\bar{x}) + \sum_{j=1}^{q} \bar{\mu}_j h_j(\bar{x}),$$

这导致与 $\bar{x}$ 是(P)的最优解相矛盾．▯

定理 3.4.7 称为凸规划的 Kuhn-Tucker 定理，其中的(3.4.19)称为凸规划问题(MP)的最优解的 Kuhn-Tucker 条件．

推论 3.4.8 设 $f: R^n \to R$ 和 $g_i: R^n \to R\ (i=1, \cdots, p)$ 是

正常凸函数，h_j: $R^n \to R$ $(j=1,\cdots,q)$ 是线性函数，$X \neq \varnothing$ 由 (3.4.1) 确定(其中 $\mathcal{Y} = R^n$)，并且 X 满足 Slater 约束规格，则 $\tilde{x} \in \tilde{X}$ 当且仅当存在 $(\tilde{\lambda}_1, \cdots, \tilde{\lambda}_p)^T \in R_+^p$ 和 $(\tilde{\mu}_1, \cdots, \tilde{\mu}_q)^T \in R^q$，使得

$$\begin{cases} \nabla f(\tilde{x}) + \sum\limits_{i=1}^{p} \tilde{\lambda}_i \nabla g_i(\tilde{x}) + \sum\limits_{j=1}^{q} \tilde{\mu}_j \nabla h_j(\tilde{x}) = 0; \\ \tilde{\lambda}_i g_i(\tilde{x}) = 0,\ i = 1, \cdots, p. \end{cases}$$

证明　由定理 3.4.7 和推论 3.2.5 即可得证. ▯

第 4 章　Lipschitz 函数的 G-次微分

Lipschitz 函数是一类重要的非光滑函数，它的 G-次微分是目前各种意义下次微分中最具有代表性的. Lipschitz 函数的 G-次微分具有一般可微函数的许多有用的性质，如次可加性、乘法规则、复合函数链式法则以及中值定理等. 此外，在有限维空间中的 Lipschitz 函数是几乎处处可微的，而对于凸的 Lipschitz 函数而言，则它几乎就是可微的.

本章将先介绍 Lipschitz 函数的 G-方向导数和 G-次微分概念. 然后，讨论它们的有关性质、几何特性以及在有限维空间的某些特殊结果.

§4.1　G-方向导数

设 $\mathscr{B}$ 是 Banach 空间，$\boldsymbol{x} \in \mathscr{B}$，$\|\boldsymbol{x}\|$（或 $\|\boldsymbol{x}\|_{\mathscr{B}}$）表示 $\boldsymbol{x}$（在 $\mathscr{B}$ 中）的范数，B 表示（$\mathscr{B}$ 中的）开单位球，即

$$B = \{\boldsymbol{x} \in \mathscr{B} \mid \|\boldsymbol{x}\| < 1\}.$$

定义 4.1.1　设集合 $S \subset \mathscr{B}$ 非空，$f: S \to R$ 是实值函数.

(1) 若对任意的 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in S$ 和某常数 $\gamma > 0$ 有

$$\left|f(\boldsymbol{x}^1) - f(\boldsymbol{x}^2)\right| \leqslant \gamma \|\boldsymbol{x}^1 - \boldsymbol{x}^2\|,$$

则称函数 f 在 S 上满足（γ 秩）Lipschitz 条件，γ 称为是 f 的 Lipschitz 常数.

(2) 若对点 $\boldsymbol{x}^0 \in S$，存在 $\varepsilon > 0$，使得 f 在 $\boldsymbol{x}^0 + \varepsilon B$ 上满足(γ 秩)Lipschitz 条件，则称 f 在点 $\boldsymbol{x}^0$ 附近满足(γ 秩)Lipschitz 条件，或 f 在点 $\boldsymbol{x}^0$ 处是(γ 秩)Lipschitz 的. 若对任意的 $\boldsymbol{x} \in S$，f 在点 $\boldsymbol{x}$ 附近都满足(γ 秩)Lipschitz 条件，则称 f 是 S 上的(γ 秩)Lipschitz 函数(泛函)，或 f 在 S 上是(γ 秩)Lipschitz 的.

Lipschitz 函数类包含了很大一类函数. 例如，凸函数就是其有效域上的 Lipschitz 函数，可微函数都是 Lipschitz 函数.

定义 4.1.2　设集合 $S \subset \mathscr{B}$ 非空，函数 $f: S \to R$ 在点 $\boldsymbol{x}^0 \in S$ 处是(γ 秩)Lipschitz 的，$\boldsymbol{d} \in \mathscr{B}$. 若极限

$$f^0(\boldsymbol{x}^0; \boldsymbol{d}) = \limsup_{\substack{\boldsymbol{x} \to \boldsymbol{x}^0 \\ t \to 0^+}} \frac{f(\boldsymbol{x} + t\boldsymbol{d}) - f(\boldsymbol{x})}{t} \tag{4.1.1}$$

存在，则称它是 f 在点 $\boldsymbol{x}^0$ 处沿方向 $\boldsymbol{d}$ 的广义方向导数或 G-方向导数.

定理 4.1.1　设集合 $S \subset \mathscr{B}$ 非空，f 在 $\boldsymbol{x} \in S$ 处是(γ 秩)Lipschitz 的.

(1) $f^0(\boldsymbol{x}; \boldsymbol{d})$关于 $\boldsymbol{d}$ 是正齐次和次可加的(即次线性的)，有限的，并且满足

$$\left| f^0(\boldsymbol{x}; \boldsymbol{d}) \right| \leqslant \gamma \| \boldsymbol{d} \|.$$

(2) $f^0(\boldsymbol{x}; -\boldsymbol{d}) = -f^0(\boldsymbol{x}; \boldsymbol{d})$.

证明　(1) 由 f 在 $\boldsymbol{x}$ 附近满足 γ 秩 Lipschitz 条件，根据定义 4.1.2 可知，对任意的 $\alpha \geqslant 0$ 有

$$f^0(\boldsymbol{x}; \alpha\boldsymbol{d}) = \limsup_{\substack{\boldsymbol{x}' \to \boldsymbol{x} \\ t \to 0^+}} \frac{f(\boldsymbol{x}' + \alpha t\boldsymbol{d}) - f(\boldsymbol{x}')}{t} = \alpha f^0(\boldsymbol{x}; \boldsymbol{d}),$$

故正齐次性成立. 又对任何 $\boldsymbol{d}^1, \boldsymbol{d}^2 \in \mathscr{B}$ 有

$$f^0(\boldsymbol{x}; \boldsymbol{d}^1 + \boldsymbol{d}^2) = \limsup_{\substack{\boldsymbol{x}' \to \boldsymbol{x} \\ t \to 0^+}} \frac{f(\boldsymbol{x}' + t\boldsymbol{d}^1 + t\boldsymbol{d}^2) - f(\boldsymbol{x}')}{t}$$

$$\leqslant \limsup_{\substack{x'+td^2\to x\\ t\to 0^+}} \frac{f(x'+td^1+td^2)-f(x'+td^2)}{t}$$

$$+\limsup_{\substack{x'\to x\\ t\to 0^+}} \frac{f(x'+td^2)-f(x')}{t},$$

即

$$f^0(x;\, d^1+d^2)\leqslant f^0(x;\, d^1)+f^0(x;\, d^2),$$

故它具有次可加性. 由此,$f^0(x;\, d)$关于 d 是次线性泛函.

另外,按定义 4.1.2 和定义 4.1.1 的(1),有

$$\left|f^0(x;\, d)\right|\leqslant \frac{\left|f(x'+td)-f(x')\right|}{t}\leqslant \gamma\|d\|,$$

因此,对于任给定的 d, $\left|f^0(x;\, d)\right|$ 是有限的.

(2) 由定义 4.1.2 知

$$f^0(x;\, -d)=\limsup_{\substack{x'\to x\\ t\to 0^+}} \frac{f(x'-td)-f(x')}{t},$$

令 $x''=x'-td$, 从上式有

$$f^0(x;\, -d)=-\limsup_{\substack{x''\to x\\ t\to 0^+}} \frac{f(x''+td)-f(x'')}{t},$$

故 $f^0(x;\, -d)=-f^0(x;\, d)$. ▯

注 4.1.1　由定理 4.1.1 的(1)和(2),易知有

$$f^0(x;\, \alpha d)=\alpha f^0(x;\, d),\ \alpha\in R. \tag{4.1.2}$$

定理 4.1.2　设集合 $S\subset\mathscr{B}$ 非空,f 在 $x\in S$ 处是(γ 秩)Lipschitz 的.

(1) 函数 $f^0(x;\, d)$在点$(x,\, d)$处是上半连续的.

(2) $f^0(x;\, d)$关于 d 在 $\mathscr{B}$ 上是 γ 秩 Lipschitz 的.

证明　(1) 考虑任意点列$\{x_k\}$, $x_k\to x$, 以及 $\{d^k\}\subset\mathscr{B}$, $d^k\to d$ $(k\to\infty)$. 对于每个 k,由定义 4.1.2 知存在 $x'_k\in\mathscr{B}$ 和 $t_k>0$

使得

$$\| x_k' - x_k \| + t_k \leqslant \frac{1}{k},$$

$$f^0(x_k;\, d^k) - \frac{1}{k} \leqslant \frac{f(x_k' + t_k d^k) - f(x_k')}{t_k}$$

$$= \frac{f(x_k' + t_k d) - f(x_k')}{t_k}$$

$$+ \frac{f(x_k' + t_k d^k) - f(x_k' + t_k d)}{t_k},$$

并且上式右端第 2 式小于 $\gamma \| d^k - d \|$. 令 $k \to \infty$,从上式即有

$$\limsup_{k \to \infty} f^0(x_k;\, d^k) \leqslant f^0(x;\, d).$$

由此,根据注 2.1.3 即知 $f^0(x;\, d)$在$(x,\, d)$处是上半连续的.

(2) 对于任意给定的 $d^1,\, d^2 \in \mathscr{B}$,我们有

$$f(x' + td^1) - f(x') \leqslant f(x' + td^2) - f(x') + \gamma \| d^1 - d^2 \| t,$$

从而

$$\limsup_{\substack{x' \to x \\ t \to 0^+}} \frac{f(x' + td^1) - f(x')}{t} \leqslant \limsup_{\substack{x' \to x \\ t \to 0^+}} \frac{f(x' + td^2) - f(x')}{t}$$

$$+ \gamma \| d^1 - d^2 \|.$$

据此,由定义 4.1.2 得

$$f^0(x;\, d^1) \leqslant f^0(x;\, d^2) + \gamma \| d^1 - d^2 \|.$$

由 $d^1,\, d^2 \in \mathscr{B}$ 的任意性,按定义 4.1.1 知 $f^0(x;\, d)$关于 d 在 $\mathscr{B}$ 上是 γ 秩 Lipschitz 的. □

§ 4.2　G-次微分

设 $\mathscr{B}$ 是 Banach 空间,$\mathscr{B}^*$ 是 $\mathscr{B}$ 的对偶空间,$x^* \in \mathscr{B}^*$ 是 $x \in$

$\mathscr{B}$ 的连续线性泛函，$\langle \boldsymbol{x}^*, \boldsymbol{x}\rangle$表示泛函 $\boldsymbol{x}^*$ 在 $\boldsymbol{x}$ 处的值.

定义 4.2.1 设集合 $S \subset \mathscr{B}$ 非空，函数 $f: S \to R$ 在点 $\boldsymbol{x}^0 \subset S$ 处是(γ 秩)Lipschitz 的. 若 $\boldsymbol{x}^* \in \mathscr{B}^*$，并且

$$\langle \boldsymbol{x}^*, \boldsymbol{d}\rangle \leqslant f^0(\boldsymbol{x}^0; \boldsymbol{d}) \quad \forall \boldsymbol{d} \in \mathscr{B}, \tag{4.2.1}$$

则称 $\boldsymbol{x}^*$ 是 f 在点 $\boldsymbol{x}^0$ 处的广义次梯度或 G-次梯度. f 在点 $\boldsymbol{x}^0$ 处的所有 G-次梯度的集合称为 f 在点 $\boldsymbol{x}^0$ 处的广义次微分或 G-次微分，记作 $\partial^0 f(\boldsymbol{x}^0)$，即

$$\partial^0 f(\boldsymbol{x}^0) = \{\boldsymbol{x}^* \in \mathscr{B}^* \mid \langle \boldsymbol{x}^*, \boldsymbol{d}\rangle \leqslant f^0(\boldsymbol{x}^0; \boldsymbol{d}) \quad \forall \boldsymbol{d} \in \mathscr{B}\}. \tag{4.2.2}$$

若 $\partial^0 f(\boldsymbol{x}^0) \neq \varnothing$，则称 f 在点 $\boldsymbol{x}^0$ 处是广义次可微的或 G-次可微的，若 f 在 S 的每一点处是 G-次可微的，则称 f 在 S 上是 G-次可微的.

例 4.2.1 设 $\mathscr{B} = R$, $f(x) = |x|$. 按定义 4.1.2，对任意的 $d \in R$ 有 $f^0(x; d) = d$. 由定义 4.2.1，得

$$\partial^0 f(x) = \begin{cases} 1, & x > 0; \\ [-1, 1], & x = 0; \\ -1, & x < 0. \end{cases}$$

定理 4.2.1 设集合 $S \subset \mathscr{B}$ 非空，f 在 $\boldsymbol{x} \in S$ 处是 γ 秩 Lipschitz 的.

(1) $\partial^0 f(\boldsymbol{x})$是非空凸集，且在 $\mathscr{B}^*$ 中是紧的(在弱 * 意义下).

(2) 对任意的 $\boldsymbol{d} \in \mathscr{B}$，有

$$f^0(\boldsymbol{x}; \boldsymbol{d}) = \max\{\langle \boldsymbol{x}^*, \boldsymbol{d}\rangle \mid \boldsymbol{x}^* \in \partial^0 f(\boldsymbol{x})\}.$$

证明 (1) 由定理 4.1.1 和 Hahn-Banach 定理[21]，至少存在一线性泛函 $\boldsymbol{x}^*: \mathscr{B} \to R$，使得对任意的 $\boldsymbol{d} \in \mathscr{B}$ 有 $\langle \boldsymbol{x}^*, \boldsymbol{d}\rangle \leqslant f^0(\boldsymbol{x}; \boldsymbol{d})$. 因此，$\partial^0 f(\boldsymbol{x})$是非空的. $\partial^0 f(\boldsymbol{x})$显然是凸的. 此外，利用定理 4.1.1 的(1)，有

$$\langle x^*, d\rangle \leqslant f^0(x; d) \leqslant \gamma \| d \|.$$

由 Alaoglu 定理[20]即知在弱 * 意义下 $\partial^0 f(x)$ 是紧的.

(2) 反证:假设对某个 $d^0 \in \mathscr{B}$, 有

$$f^0(x; d^0) > \langle x^*, d^0\rangle \quad \forall x^* \in \partial^0 f(x).$$

由 Hahn-Banach 定理可知,存在连续线性泛函 $\bar{x}^* \in \mathscr{B}^*$, 使得

$$f^0(x; d) \geqslant \langle \bar{x}^*, d\rangle \quad \forall d \in \mathscr{B},$$

并且 $f^0(x; d^0) = \langle \bar{x}^*, d^0\rangle$. 然而 $\bar{x}^* \in \partial^0 f(x)$, 这显然是矛盾的. ▯

定理 4.2.2　设集合 $S \subset \mathscr{B}$ 非空, f 在 $x \in S$ 处是 Lipschitz 的.

(1) $x^* \in \partial^0 f(x)$ 当且仅当

$$\langle x^*, d\rangle \leqslant f^0(x; d) \quad \forall d \in \mathscr{B}.$$

(2) 设 $\{x_k\} \subset \mathscr{B}$, $\{x_k^*\} \subset \mathscr{B}^*$, $x_k^* \in \partial^0 f(x_k)$ $(k = 1, 2, \cdots)$. 若 $x_k \to x$, $x_k^* \to x^*$ $(k \to \infty)$, 则 $x^* \in \partial^0 f(x)$.

(3) 设 $B \subset \mathscr{B}$ 是开单位球,则

$$\partial^0 f(x) = \bigcap_{\delta > 0} \bigcup_{x' \in x + \delta B} \partial^0 f(x').$$

(4) 若 $B \subset R^n$, 则对任给的 $\varepsilon > 0$, 存在 $\delta > 0$ 使得

$$\partial f^0(x') \subset \partial f^0(x) + \varepsilon B \quad \forall x' \in x + \delta B.$$

证明　(1) 由定义 4.2.1 可得.

(2) 利用定理 4.1.2 的(1)和 $\langle x_k^*, d\rangle \leqslant f^0(x_k; d)$, 令 $k \to \infty$, 有

$$\langle x^*, d\rangle \leqslant f^0(x; d) \quad \forall d \in \mathscr{B}.$$

因此,得知 $x^* \in \partial^0 f(x)$.

(3)和(4)均可由(2)推得. ▯

定理 4.2.3　设 $f: \mathscr{B} \to R$ 在 $\bar{x} \in \mathscr{B}$ 处是 Lipschitz 的. 若 $\bar{x}$ 是

无约束极小化问题(UMP)的最优解,则 $\mathbf{0}\in\partial^{\circ}f(\bar{x})$.

证明 由定义 3.4.1 和定义 4.1.2 可知

$$f^{\circ}(\bar{x};\boldsymbol{d})\geqslant 0\quad \forall \boldsymbol{d}\in\mathscr{B}.$$

由此,按定义 4.2.1 得 $\mathbf{0}\in\partial^{\circ}f(\bar{x})$. ∎

下面讨论 G-次微分与其他意义下的导数的关系.

设 $\mathscr{D}$ 是 Banach 空间.

定义 4.2.2 设 $\varphi:\mathscr{B}\to\mathscr{D}$ 是映射,点 $\boldsymbol{x}^0\in\mathscr{B}$. 若对 $\boldsymbol{x}^0$ 附近的任意点 $\boldsymbol{x}^1$、$\boldsymbol{x}^2$ 和某常数 $\gamma>0$ 有

$$\left\|\varphi(\boldsymbol{x}^1)-\varphi(\boldsymbol{x}^2)\right\|_{\mathscr{D}}\leqslant\gamma\|\boldsymbol{x}^1-\boldsymbol{x}^2\|_{\mathscr{B}},$$

则称映射 φ 在点 $\boldsymbol{x}^0$ 附近满足(γ 秩)Lipschitz 条件,或在点 $\boldsymbol{x}^0$ 处是 Lipschitz 的.

设 $L(\mathscr{B},\mathscr{D})$表示 $\mathscr{B}$ 到 $\mathscr{D}$ 的所有连续线性映射组成的集合.

定义 4.2.3 设 $S\subset\mathscr{B}$ 是紧集,$\varphi:S\to\mathscr{D}$ 是映射,点 $\boldsymbol{x}^0\in\mathscr{B}$,$D_S\varphi(\boldsymbol{x}^0)\in L(\mathscr{B},\mathscr{D})$. 若对任意的 $\boldsymbol{d}\in\mathscr{B}$,有

$$\lim_{\substack{\boldsymbol{x}\to\boldsymbol{x}^0\\ t\to 0^+}}\frac{\varphi(\boldsymbol{x}+t\boldsymbol{d})-\varphi(\boldsymbol{x})}{t}=\langle D_S\varphi(\boldsymbol{x}^0),\boldsymbol{d}\rangle \tag{4.2.3}$$

成立,并且左端的极限关于 $\boldsymbol{d}$ 一致收敛,则称 φ 在点 $\boldsymbol{x}^0$ 处是严格可微的,并且称 $D_S\varphi(\boldsymbol{x}^0)$是 φ 在点 $\boldsymbol{x}^0$ 处的严格导数. 若 φ 在 S 上的每一点处是严格可微的,则称 φ 是 S 上的严格可微映射,或在 S 上是严格可微的.

定理 4.2.4 设 $S\subset\mathscr{B}$ 是紧集,$\varphi:\mathscr{B}\to\mathscr{D}$ 是映射,$\boldsymbol{x}^*\in L(\mathscr{B},\mathscr{D})$. φ 在 $\boldsymbol{x}\in S$ 处是严格可微的且 $D_S\varphi(\boldsymbol{x})=\boldsymbol{x}^*$ 的充要条件是:φ 在 $\boldsymbol{x}$ 附近满足 Lipschitz 条件,并且对任意的 $\boldsymbol{d}\in\mathscr{B}$ 有

$$\lim_{\substack{\boldsymbol{x}'\to\boldsymbol{x}\\ t\to 0^+}}\frac{\varphi(\boldsymbol{x}'+t\boldsymbol{d})-\varphi(\boldsymbol{x}')}{t}=\langle\boldsymbol{x}^*,\boldsymbol{d}\rangle. \tag{4.2.4}$$

证明 必要性. 由 φ 在 $\boldsymbol{x}$ 处是严格可微的及 $D_S\varphi(\boldsymbol{x})=\boldsymbol{x}^*$,按定义 4.2.3 可知,对任意的 $\boldsymbol{d}\in\mathscr{B}$,(4.2.4)成立. 下面证明 φ 在 $\boldsymbol{x}$

附近满足 Lipschitz 条件. 反证：假设不然，则存在点列 $\{x_k\}$, $x_k \to x$, $\{x'_k\}$, $x'_k \to x'$ ($k \to \infty$)，使得对每个 k 有 $x_k, x'_k \in x + (1/k)B$ 和

$$\left\|\varphi(x'_k) - \varphi(x_k)\right\|_{\mathscr{D}} \geqslant k \| x'_k - x_k \|_{\mathscr{B}}.$$

再定义序列 $\{t_k\}$, $t_k > 0$, 以及 $\{d^k\}$ 使得 $x'_k = x_k + t_k d^k$ 和 $\| d^k \| = k^{-\frac{1}{2}}$, $t_k \to 0 (k \to \infty)$. 取集合 $A = \{d^k\} \cup (\mathbf{0})$，显然它是紧集. 因此，根据定义 4.2.3，对任给的 $\varepsilon > 0$，存在 N 使得对任何 $k \geqslant N$ 和所有的 $d \in A$ 有

$$\left\|\frac{\varphi(x'_k + t_k d) - \varphi(x'_k)}{t_k} - \langle D_S\varphi(x), d \rangle\right\|_{\mathscr{D}} \leqslant \varepsilon. \quad (4.2.5)$$

但是，另一方面有

$$\left\|\frac{\varphi(x'_k + t_k d) - \varphi(x'_k)}{t_k} - \langle D_S\varphi(x), d \rangle\right\|_{\mathscr{D}} > k^{\frac{1}{2}},$$

令 $k \to \infty$，这导致与(4.2.5)相矛盾.

充分性. 我们证明 φ 在 x 处是严格可微的. 设 $S \subset \mathscr{B}$ 是任一紧子集，$\varepsilon > 0$ 是任给的数. 对任意的 $d \in \mathscr{B}$ 存在正数 $\delta(d)$，使得对任意的 $x' \in x + \varepsilon B$ 和 $t \in (0, \delta)$ 有

$$\left\|\frac{\varphi(x' + td) - \varphi(x')}{t} - \langle D_S\varphi(x), d \rangle\right\|_{\mathscr{D}} < \varepsilon. \quad (4.2.6)$$

因为 φ 在 x 处是 Lipschitz 的，故存在 δ' 使得对任意的 $d' \in d + \delta' B$ 和 $t \in (0, \delta')$ 有 $x' + td' \in x + \delta' B$，以及

$$\left\|\varphi(x' + td') - \varphi(x' + td)\right\|_{\mathscr{D}} < \gamma t \| d' - d \|_{\mathscr{B}}. \quad (4.2.7)$$

由(4.2.6)和(4.2.7)可重新选择 $\delta(d)$，使得对任意的 $d' \in d + \delta B$ 及任意的 $x' \in x + \delta B$ 和 $t \in (0, \delta)$，有

$$\left\|\frac{\varphi(x' + td') - \varphi(x')}{t} - \langle D_S\varphi(x), d \rangle\right\|_{\mathscr{D}} < 2\varepsilon.$$

由此，得到 $\{d+\delta(d)B \mid d\in\mathscr{B}\}$ 是 S 的一个开覆盖. 由于 S 是紧的，则它存在一个有限子覆盖 $d_1+\delta(d_1)B,\cdots,d_n+\delta(d_n)B$. 选择 $\delta=\min\limits_{1\leqslant k\leqslant n}\delta(d_k)$，则对任意的 $d\in S$ 及任意的 $x'\in x+\delta' B$ 和 $t\in(0,\delta')$，(4.2.6)成立. 因此，由定义 4.2.3，φ 在 x 处是严格可微的. ▯

以下考虑 G-次微分与 Frechet 导数、Gâteaux 导数以及严格导数之间的关系.

定理 4.2.5 设集合 $S\subset\mathscr{B}$ 非空，$f: S\to R$ 在 $x\in S$ 处是 Lipschitz 的，并且 f 在 x 处是 Frechet 可微的.

(1) 设 $f'(x)$ 是 f 在 x 处的 Frechet 导数，则 $f'(x)\in\partial^0 f(x)$.

(2) 设 f'_x 是 f 在 x 处的 Gâteaux 导数，则 $f'_x\in\partial^0 f(x)$.

证明 (1) 因为 f 在 x 处是 Frechet 可微的，由定理 3.1.5 和注 3.1.1 得知

$$\langle f'(x),d\rangle=f'(x;d)\quad\forall d\in\mathscr{B}.$$

又由定义 3.1.1 的(3)和定义 4.1.2，有

$$f'(x;d)\leqslant f^0(x;d)\quad\forall d\in\mathscr{B},$$

所以

$$\langle f'(x),d\rangle\leqslant f^0(x;d)\quad\forall d\in\mathscr{B}.$$

据此，按定义 4.2.1 即知 $f'(x)\in\partial^0 f(x)$.

(2) 由定理 3.1.5 知 $f'(x)=f'_x$，根据(1)即得证. ▯

定理 4.2.6 设集合 $S\subset\mathscr{B}$ 非空，则 $f: S\to R$ 在 $x\in S$ 处是严格可微的当且仅当 f 在 x 处是 Lipschitz 的，并且 $\partial^0 f(x)$ 是单点集，即 $\partial^0 f(x)=\{D_S f(x)\}$.

证明 必要性. 由于 f 在 x 处是严格可微的，从定理 4.2.4 知 f 在 x 处是 Lipschitz 的. 据此，按定义 4.1.2 和定义 4.2.3 知，对任意的 $d\in\mathscr{B}$ 有 $f^0(x;d)=\langle D_S f(x),d\rangle$. 再由定理 4.2.1 的(2)即得 $\partial^0 f(x)=\{D_S f(x)\}$.

充分性. 根据定理 4.2.1 的(2)，只要证明对任意的 $d\in\mathscr{B}$ 有

(4.2.4)成立 ($\varphi = f$)，即可推出 f 在 $\boldsymbol{x}$ 处是严格可微的. 现在证明对任意的 $\boldsymbol{d} \in \mathscr{B}$ 有

$$f^0(\boldsymbol{x};\ \boldsymbol{d}) = \langle \boldsymbol{x}^*,\ \boldsymbol{d} \rangle, \tag{4.2.8}$$

其中 $\partial^0 f(\boldsymbol{x}) = \{\boldsymbol{x}^*\}$. 反证：假设(4.2.8)不成立，则存在 $\bar{\boldsymbol{d}} \in \mathscr{B}$ 使得 $f^0(\boldsymbol{x};\ \bar{\boldsymbol{d}}) > \langle \boldsymbol{x}^*,\ \bar{\boldsymbol{d}} \rangle$. 由 Hahn-Banach 定理可知，存在一点 $\bar{\boldsymbol{x}}^* \in \mathscr{B}^*$，使得 $f^0(\boldsymbol{x};\ \boldsymbol{d}) \geqslant \langle \bar{\boldsymbol{x}}^*,\ \boldsymbol{d} \rangle$ 对任何 $\boldsymbol{d} \in \mathscr{B}$ 成立. 特别是，有 $f^0(\boldsymbol{x};\ \bar{\boldsymbol{d}}) = \langle \bar{\boldsymbol{x}}^*;\ \bar{\boldsymbol{d}} \rangle$，故 $\bar{\boldsymbol{x}}^* \in \partial^0 f(\boldsymbol{x})$. 因为已知 $\partial^0 f(\boldsymbol{x})$是单点集，所以 $\bar{\boldsymbol{x}}^* = \boldsymbol{x}^*$. 从而导致 $f^0(\boldsymbol{x};\ \bar{\boldsymbol{d}}) > f^0(\boldsymbol{x};\ \bar{\boldsymbol{d}})$，产生矛盾. 从(4.2.8)，我们有

$$\begin{aligned}
&\liminf_{\substack{x' \to x \\ t \to 0^+}} \frac{f(\boldsymbol{x}' + t\boldsymbol{d}) - f(\boldsymbol{x}')}{t} \\
&= -\limsup_{\substack{x' \to x \\ t \to 0^+}} \frac{f(\boldsymbol{x}') - f(\boldsymbol{x}' + t\boldsymbol{d})}{t} \\
&= -\limsup_{\substack{x' \to x \\ t \to 0^+}} \frac{f(\boldsymbol{x}' + t\boldsymbol{d} - t\boldsymbol{d}) - f(\boldsymbol{x}' + t\boldsymbol{d})}{t} \\
&= -f^0(\boldsymbol{x};\ -\boldsymbol{d}) = -\langle \bar{\boldsymbol{x}}^*,\ -\boldsymbol{d} \rangle \\
&= \langle \boldsymbol{x}^*,\ \boldsymbol{d} \rangle = f^0(\boldsymbol{x};\ \boldsymbol{d}) \\
&= -\limsup_{\substack{x' \to x \\ t \to 0^+}} \frac{f(\boldsymbol{x}' + t\boldsymbol{d}) - f(\boldsymbol{x}')}{t}.
\end{aligned}$$

于是，对于任意的 $\boldsymbol{d} \in \mathscr{B}$，(4.2.4)成立 ($\varphi = f$). □

推论 4.2.7　设 $f: R^n \to R$ 在 $\boldsymbol{x} \in R^n$ 处是 Lipschitz 的，$\varepsilon > 0$，则对任意的 $\boldsymbol{x}' \in \boldsymbol{x} + \varepsilon B$，$\partial^0 f(\boldsymbol{x}')$是单点集当且仅当 f 在 $\boldsymbol{x} + \varepsilon B$ 上是连续可微的.

证明　由定理 4.2.6 可得证. □

关于 Lipschitz 函数的 G-次微分与凸函数的次微分之间有以下关系.

定理 4.2.8 设 $S\subset\mathscr{B}$ 是开凸集，$f: S\to R$ 是凸函数. 若 f 在 S 中一点的某邻域有界，则对任意的 $\boldsymbol{x}\in S$，f 在 $\boldsymbol{x}$ 处是 Lipschitz 的.

证明 首先证明对任意的 $\boldsymbol{x}\in S$，f 在 $\boldsymbol{x}$ 的一个邻域有界. 不妨设 f 在点 $\boldsymbol{0}$ 的邻域 $\varepsilon B\subset S$ 上是有界的，即有 $M\in R$ 使

$$f(\boldsymbol{x}')\leqslant M\quad\forall\boldsymbol{x}'\in\varepsilon B. \tag{4.2.9}$$

由于 S 是开凸集，可以选择 $\rho>1$ 使 $\boldsymbol{y}=\rho\boldsymbol{x}\in S$. 令 $\lambda=1/\rho$，作集合

$$D=\{\boldsymbol{d}\in\mathscr{B}\,|\,\boldsymbol{d}=(1-\lambda)\boldsymbol{x}'+\lambda\boldsymbol{y},\ \boldsymbol{x}'\in\varepsilon B\},$$

则 $D=\boldsymbol{x}+(1-\lambda)\varepsilon B$. 对任意的 $\boldsymbol{d}\in D$，由(4.2.9)和 f 是 S 上的凸函数，有

$$f(\boldsymbol{d})\leqslant(1-\lambda)f(\boldsymbol{x}')+\lambda f(\boldsymbol{y})\leqslant M+\lambda f(\boldsymbol{y}). \tag{4.2.10}$$

取 $\boldsymbol{d}'\in D$ 使得 $\boldsymbol{x}=\dfrac{\boldsymbol{d}+\boldsymbol{d}'}{2}$，则有

$$f(\boldsymbol{x})\leqslant\frac{1}{2}f(\boldsymbol{d})+\frac{1}{2}f(\boldsymbol{d}').$$

由(4.2.10)得

$$f(\boldsymbol{d})\geqslant 2f(\boldsymbol{x})-f(\boldsymbol{d}')\geqslant 2f(\boldsymbol{x})-M-\lambda f(\boldsymbol{y}). \tag{4.2.11}$$

从(4.2.10)和(4.2.11)知 f 在 $\boldsymbol{x}$ 的邻域有界，取

$$N=\begin{cases}M+\lambda(\boldsymbol{y}), & f(\boldsymbol{x})\geqslant 0;\\ M+\lambda f(\boldsymbol{y})-2f(\boldsymbol{x}), & f(\boldsymbol{x})<0,\end{cases}$$

故存在 $\delta>0$，使得

$$|f(\boldsymbol{v})|\leqslant N\quad\forall\boldsymbol{v}\in\boldsymbol{x}+2\delta B. \tag{4.2.12}$$

对任意的 $\boldsymbol{x}^1$，$\boldsymbol{x}^2 \in \boldsymbol{x} + 2\delta B$，设 $\boldsymbol{x}^3 = \boldsymbol{x}^2 + \dfrac{\delta}{\alpha}(\boldsymbol{x}^2 - \boldsymbol{x}^1)$，其中 $\alpha = \| \boldsymbol{x}^2 - \boldsymbol{x}^1 \|$，显然 $\boldsymbol{x}^3 \in \boldsymbol{x} + 2\delta B$. 因此，我们有

$$\boldsymbol{x}^2 = \frac{\delta}{\delta + \alpha}\boldsymbol{x}^1 + \frac{\alpha}{\delta + \alpha}\boldsymbol{x}^3.$$

因为 f 在 D 上是凸的，所以

$$f(\boldsymbol{x}^2) \leqslant \frac{\delta}{\delta + \alpha}f(\boldsymbol{x}^1) + \frac{\alpha}{\delta + \alpha}f(\boldsymbol{x}^3),$$

从而

$$\begin{aligned} f(\boldsymbol{x}^2) - f(\boldsymbol{x}^1) &\leqslant \frac{\delta}{\delta + \alpha}\left[f(\boldsymbol{x}^3) - f(\boldsymbol{x}^1)\right] \\ &\leqslant \frac{\alpha}{\delta}\left[f(\boldsymbol{x}^3) - f(\boldsymbol{x}^1)\right]. \end{aligned}$$

由(4.2.12)则得到

$$f(\boldsymbol{x}^2) - f(\boldsymbol{x}^1) \leqslant \frac{2N}{\delta}\| \boldsymbol{x}^2 - \boldsymbol{x}^1 \|.$$

注意 $\boldsymbol{x}^1$ 和 $\boldsymbol{x}^2$ 是 $\boldsymbol{x} + 2\delta B$ 中的任意点，因此得知 f 在 $\boldsymbol{x}$ 处是 Lipschitz 的. ▯

定理 4.2.9　设 $S \subset \mathscr{B}$ 是开凸集. 若 $f: S \to R$ 是凸函数，并且 f 在 $\boldsymbol{x} \in S$ 处是 Lipschitz 的，则

$$f^0(\boldsymbol{x}; \boldsymbol{d}) = f'(\boldsymbol{x}; \boldsymbol{d}) \quad \forall d \in \mathscr{B}.$$

证明　对任给定的 $\delta > 0$，我们有

$$f^0(\boldsymbol{x}; \boldsymbol{d}) = \lim_{\varepsilon \to 0^+} \sup_{\|\boldsymbol{x}' - \boldsymbol{x}\| \leqslant \varepsilon\delta} \sup_{0 < t < \varepsilon} \frac{f(\boldsymbol{x}' + t\boldsymbol{d}) - f(\boldsymbol{x}')}{t},$$

由 f 是凸函数可知，映射：$t \to \dfrac{f(\boldsymbol{x}' + t\boldsymbol{d}) + f(\boldsymbol{x}')}{t}$ 关于 t 是非减函数. 因此，有

$$f^0(x;d)=\lim_{\varepsilon\to0^+}\sup_{\|x'-x\|\leqslant\varepsilon\delta}\frac{f(x'+\varepsilon d)-f(x')}{\varepsilon}.$$

再根据 f 在 $x'\in x+\varepsilon\delta B$ 上是 Lipschitz 的，可知有 $\gamma>0$ 使得

$$\left|\frac{f(x'+\varepsilon d)-f(x')}{\varepsilon}-\frac{f(x+\varepsilon d)-f(x)}{\varepsilon}\right|\leqslant 2\gamma\delta,$$

故

$$\begin{aligned}f^0(x;d)&\leqslant\lim_{\varepsilon\to0^+}\frac{f(x+\varepsilon d)-f(x)}{\varepsilon}+2\delta\gamma\\&=f'(x;d)+2\gamma\delta.\end{aligned}$$

由 δ 的任意性推知 $f^0(x;d)\leqslant f'(x;d)$，又显然有 $f'(x;d)\leqslant f^0(x;d)$，因而 $f^0(x;d)=f'(x;d)$. ▯

定理 4.2.9 表明，对于在某点附近满足 Lipschitz 条件的凸函数，它的次微分与 G-次微分是一致的.

定理 4.2.10 设集合 $S\subset\mathscr{B}$ 非空，f: $S\to R$ 在 $x\in S$ 处是 Lipschitz 的，则 $f^0(x;d)$ 是 $\partial^0 f(x)$的支撑函数.

证明 由定义 4.2.1 和定义 1.3.4 即得证. ▯

设集合 $T\subset\mathscr{B}^*$ 非空，记 $\sigma_T^*:\mathscr{B}\to R\cup\{+\infty\}$ 是 T 的支撑函数.

定理 4.2.11 设 $S_1,S_2\subset\mathscr{B}$ 是非空闭凸集，$T_1,T_2\subset\mathscr{B}^*$ 是非空弱闭凸集.

(1) $S_1\subset S_2$ 当且仅当 $\sigma_{S_1}(x^*)\leqslant\sigma_{S_2}(x^*)\quad(\forall x^*\in\mathscr{B}^*)$.

(2) $T_1\subset T_2$ 当且仅当 $\sigma_{T_1}^*(x)\leqslant\sigma_{T_2}^*(x)\quad(\forall x\in\mathscr{B})$.

(3) T_1 是弱 * 紧的当且仅当 $\sigma_{T_1}^*(x)$在 $\mathscr{B}$ 上取有限值.

(4) 函数 σ: $\mathscr{B}\to R\cup\{+\infty\}$ 是正齐次和次可加的，下半连续的，以及不恒为$+\infty$的充要条件是：存在非空弱闭凸子集 $T\subset\mathscr{B}^*$ 使得 $\sigma=\sigma_T^*$，并且 T 是唯一的.

证明 (1)和(2)由定义 1.3.2 可得证.

(3) 由 T_1 是弱 * 紧的，对任意的 $x\in\mathscr{B}$，从 $x^*\in T_1$ 知$\langle x^*,x\rangle$为有限值. 按定义 1.3.4 得 $\sigma_{T_1}^*(x)$在 $\mathscr{B}$ 上取有限值. 反之，显然

成立.

(4) 充分性显然成立,以下证明必要性. 设集合

$$T = \{x^* \in \mathscr{B}^* \mid \sigma(x) \geqslant \langle x^*, x\rangle, x \in \mathscr{B}\}.$$

首先证明 T 是凸集. 设对任意的 $x_1^*, x_2^* \in T$,任意的 $\lambda \in (0, 1)$,有

$$\sigma(x) \geqslant \langle x_1^*, x\rangle \quad \forall x \in \mathscr{B},$$

$$\sigma(x) \geqslant \langle x_2^*, x\rangle \quad \forall x \in \mathscr{B}.$$

因为 σ 是正齐次和次可加的,则对任意的 $x \in \mathscr{B}$, 有

$$\begin{aligned}\sigma(x) &= \sigma(\lambda x + (1-\lambda)x)\\ &\geqslant \lambda\sigma(x) + (1-\lambda)\sigma(x)\\ &\geqslant \langle \lambda x_1^* + (1-\lambda)x_2^*, x\rangle.\end{aligned}$$

由 T 的定义得 $\lambda x_1^* + (1-\lambda)x_2^* \in T$, 故 T 是凸集. 为证明 T 是弱闭的,设点列 $\{x_k^*\} \subset T$ $(k = 1, 2, \cdots)$,对任意的 $x \in \mathscr{B}$ 有 $\langle x_k^*, x\rangle \to \langle x^*, x\rangle$(弱收敛)$(k\to\infty)$. 从

$$\sigma(x) \geqslant \langle x_k^*, x\rangle \quad \forall x \in \mathscr{B},\ k = 1, 2, \cdots,$$

得

$$\sigma(x) \geqslant \langle x^*, x\rangle \quad \forall x \in \mathscr{B},$$

则有 $x^* \in T$. T 的非空性的证明类似于定理 4.2.1 的(1)的证明. 下面证明

$$\sigma(x) = \sup_{x^* \in T} \langle x^*, x\rangle. \tag{4.2.13}$$

反证:假设对某个 $x^0 \in \mathscr{B}$, 有

$$\sigma(x^0) > \langle x^*, x^0\rangle \quad \forall x^* \in T.$$

根据 Hahn-Banach 定理,存在一连续线性泛函 $\overline{x}^* \in \mathscr{B}^*$,使得

$$\sigma(x) \geqslant \langle \overline{x}^*, x\rangle \quad \forall x \in \mathscr{B},$$

并且 $\sigma(x^0)=\langle\bar{x}^*,x^0\rangle$. 然而 $\bar{x}^*\in T$. 这显然是矛盾的. 从而(4.2.13)得证. ▯

最后,介绍函数在一点处正则的概念.

定义 4.2.4 设 $f:\mathscr{B}\to R$ 是实值函数,点 $x^0\in\mathscr{B}$. 若对任意的 $d\in\mathscr{B}$, f 在点 x^0 处沿 d 的右方向导数 $f'_+(x^0;d)$ 存在,并且对任意的 $d\in\mathscr{B}$,有 $f^0(x^0;d)=f'_+(x^0;d)$,则称函数 f 在点 x^0 处是正则的.

定理 4.2.12 设 $f:\mathscr{B}\to R$ 在 $x\in\mathscr{B}$ 处是 Lipschitz 的.

(1) 若 f 在 x 处是严格可微的,则 f 在 x 处是正则的.

(2) 若 f 是凸函数,则 f 在 x 处是正则的.

(3) 若 f 在 x 处存在 Gâteaux 导数 f'_x,并且 f 在 x 处是正则的,则 $\partial^0 f(x)=\{f'_x\}$.

证明 (1) 由定义 4.2.3、定理 4.2.2 和定理 4.2.6,有

$$f^0(x;d)=\max\{\langle x^*,d\rangle\mid x^*\in\partial^0 f(x)\}$$

$$=\langle D_S f(x),d\rangle=f'(x;d).$$

因此,按定义 4.2.4 即得 f 在 x 处是正则的.

(2) 由定理 4.2.9 可知,对任意的 $d\in\mathscr{B}$ 有 $f'(x;d)=f^0(x;d)$.

(3) 由定义 4.2.4 可以推得. ▯

定理 4.2.13 设 $f_1,f_2:\mathscr{B}\to R$ 在 $x\in\mathscr{B}$ 处是正则的, $\alpha>0$, $\beta>0$, 则 $\alpha f_1+\beta f_2$ 在 x 处是正则的.

证明 因为 $f_1,f_2:\mathscr{B}\to R$ 在 x 处是正则的, $\alpha>0$, $\beta>0$, 故对任意的 $d\in\mathscr{B}$ 有

$$(\alpha f_1+\beta f_2)^0(x;d)\geqslant(\alpha f_1+\beta f_2)'(x;d)$$

$$=\alpha f'_1(x;d)+\beta f'_2(x;d)$$

$$=\alpha f^0_1(x;d)+\beta f^0_2(x;d)$$

$$\geqslant(\alpha f_1+\beta f_2)^0(x;d).$$

据此，有

$$(\alpha f_1 + \beta f_2)^0(\boldsymbol{x};\boldsymbol{d}) = (\alpha f_1 + \beta f_2)'(\boldsymbol{x};\boldsymbol{d}) \quad \forall \boldsymbol{d} \in \mathscr{B},$$

于是由定义 4.2.4 即得证. ▯

例 4.2.2 考虑函数 $f: R^n \to R$, $\boldsymbol{x} = (x_1, \cdots, x_m)^T$,

$$f(\boldsymbol{x}) = f(x_1, \cdots, x_m) = \max\{x_i \mid i = 1, \cdots, m\}.$$

记 $I = \{i \in \{1, \cdots, m\} \mid x_i = \max\limits_{1 \leqslant j \leqslant n} x_j\}$, $\boldsymbol{d} = (d_1, \cdots, d_m)^T$，则有 $f'(\boldsymbol{x};\boldsymbol{d}) = \max\limits_{i \in I} d_i$ 和

$$\partial^0 f(\boldsymbol{0}) = \Big\{(a_1, \cdots, a_m)^T \in R^m \mid a_i \geqslant 0, i = 1, \cdots, m;$$

$$a_i = 0, i \in I, \sum_{i=1}^m a_i = 1\Big\}.$$

§ 4.3 G-次微分的性质

本节阐述 G-次微分的运算规则、中值定理以及复合函数的链式法则等性质. 我们指出，以后引进的其他意义下的次微分，都没有 G-次微分所具有的那样多好的性质.

设 $\mathscr{B}$ 是 Banach 空间. 除非有特别的申明，本节均设 $f: \mathscr{B} \to R$ 在点 $\boldsymbol{x} \in \mathscr{B}$ 附近满足 Lipschitz 条件.

定理 4.3.1 设 $\boldsymbol{x} \in \mathscr{B}$, $\alpha \in R$，则 $\partial^0(\alpha f)(\boldsymbol{x}) = \alpha\partial^0 f(\boldsymbol{x})$.

证明 当 $\alpha \geqslant 0$ 时，对任意的 $\boldsymbol{d} \in \mathscr{B}$，显然有

$$(\alpha f)^0(\boldsymbol{x};\boldsymbol{d}) = \alpha f^0(\boldsymbol{x};\boldsymbol{d}),$$

故 $\partial^0(\alpha f)(\boldsymbol{x}) = \alpha\partial^0 f(\boldsymbol{x})$. 现设 $\boldsymbol{x}^* \in \partial^0(-f)(\boldsymbol{x})$，则对任意的 $\boldsymbol{d} \in \mathscr{B}$，有 $(-f)^0(\boldsymbol{x};\boldsymbol{d}) \geqslant \langle \boldsymbol{x}^*, \boldsymbol{d} \rangle$. 由定理 4.1.1 的(2)，对任意的 $\boldsymbol{d} \in \mathscr{B}$ 有 $f^0(\boldsymbol{x};\boldsymbol{d}) \geqslant \langle -\boldsymbol{x}^*, \boldsymbol{d} \rangle$，因此得 $\boldsymbol{x}^* \in -\partial^0 f(\boldsymbol{x})$. 反之，从 $\boldsymbol{x}^* \in -\partial^0 f(\boldsymbol{x})$ 不难得知有 $\boldsymbol{x}^* \in \partial^0(-f)(\boldsymbol{x})$. ▯

定理 4.3.2 设 $f_1, f_2: \mathscr{B} \to R$ 是实值函数. 若 f_1 和 f_2 在

$x \in \mathscr{B}$ 处是 Lipschitz 的，则

$$\partial^0(f_1 + f_2)(x) \subset \partial^0 f_1(x) + \partial^0 f_2(x).$$

证明 易知 $f_1 + f_2$ 在 x 处是 Lipschitz 的，由定义 4.1.2 得

$$(f_1 + f_2)^0(x; d) = \limsup_{\substack{x' \to x \\ t \to 0^+}} \frac{f_1(x' + td) + f_2(x' + td) - f_1(x') - f_2(x')}{t}.$$

按注 2.1.3，对于任给的 $\varepsilon > 0$，存在 $\delta > 0$，使得当 $\|x' - x\| + t < \delta$ 时有

$$\begin{aligned}
&(f_1 + f_2)^0(x; d) - \varepsilon \\
&\leqslant \frac{f_1(x' + td) + f_2(x' + td) - f_1(x') - f_2(x')}{t} \\
&= \frac{f_1(x' + td) - f_1(x')}{t} + \frac{f_2(x' + td) - f_2(x')}{t}.
\end{aligned}$$

令 $x' \to x$，$t \to 0^+$，从上式知

$$(f_1 + f_2)^0(x; d) \leqslant f_1^0(x; d) + f_2^0(x; d) \quad \forall d \in \mathscr{B}.$$

设 $x^* \in \partial^0(f_1 + f_2)(x)$，由定义 4.2.1 和上式得

$$f_1^0(x; d) + f_2^0(x; d) \geqslant \langle x^*, d \rangle \quad \forall d \in \mathscr{B}.$$

据此，由定理 4.2.1 的(2)知，对任意的 $d \in \mathscr{B}$ 和任意的 $x^* \in \partial^0(f_1 + f_2)(x)$，有

$$\begin{aligned}
&\max\{\langle x_1^*, d \rangle \mid x_1^* \in \partial^0 f_1(x)\} + \max\{\langle x_2^*, d \rangle \mid x_2^* \in \partial^0 f_2(x)\} \\
&\quad \geqslant \langle x^*, d \rangle,
\end{aligned}$$

即

$$\begin{aligned}
&\max\{\langle x_1^* + x_2^*, d \rangle \mid x_1^* + x_2^* \in \partial^0 f_1(x) + \partial^0 f_2(x)\} \\
&\quad \geqslant \langle x^*, d \rangle \quad \forall d \in \mathscr{B}, \forall x^* \in \partial^0(f_1 + f_2)(x).
\end{aligned} \tag{4.3.1}$$

现假设 $x^* \notin \partial^0 f_1(x) + \partial^0 f_2(x)$，由定理 4.2.1 的(1)和定理 4.2.2 知 $\partial^0 f_1(x) + \partial^0 f_2(x)$ 是闭凸集，因此单点集 $\{x^*\}$ 与凸集 $\partial^0 f_1(x) + \partial^0 f_2(x)$ 是可严格分离的. 根据定理 1.3.7 的(2)，存在 $\bar{d} \in \mathscr{B}\setminus\{\mathbf{0}\}$，使得

$$\langle x_1^* + x_2^*, \bar{d}\rangle < \langle x^*, \bar{d}\rangle$$

$$\forall x_1^* + x_2^* \in \partial^0 f_1(x) + \partial^0 f_2(x),$$

它与(4.3.1)相矛盾，于是定理得证. □

定理 4.3.3　设 f_1 和 f_2 与定理 4.3.2 的条件相同. 若 f_1 或 f_2 在 $x \in \mathscr{B}$ 处是严格可微的，则

$$\partial^0 (f_1 + f_2)(x) = \partial^0 f_1(x) + \partial^0 f_2(x).$$

证明　不妨设 f_1 在 x 处是严格可微的，则对 $x' \in \mathscr{B}$ 和任给的 $\varepsilon > 0$，存在 $\delta > 0$，使得当 $\|x' - x\| + t < \delta$ 时有

$$\frac{f_1(x' + td) - f_1(x')}{t} - \varepsilon \leqslant f_1'(x; d)$$

$$\leqslant \frac{f_1(x' + td) - f_1(x')}{t} + \varepsilon.$$

据此，我们有

$$\frac{f_1(x' + td) + f_2(x' + td) - f_1(x') - f_2(x')}{t} - \varepsilon$$

$$\leqslant f_1'(x; d) + \frac{f_2(x' + td) - f_2(x')}{t}$$

$$\leqslant \frac{f_1(x' + td) + f_2(x' + td) - f_1(x') - f_2(x')}{t} + \varepsilon.$$

令 $x' \to x$，$t \to 0^+$，得到

$$(f_1 + f_2)^0(x; d) - \varepsilon \leqslant f_1'(x; d) + f_2^0(x; d)$$

$$\leqslant (f_1 + f_2)^0(x; d) + \varepsilon.$$

由 ϵ 的任意性，从上式有

$$(f_1+f_2)^0(\boldsymbol{x};\boldsymbol{d})=f_1^0(\boldsymbol{x};\boldsymbol{d})+f_2^0(\boldsymbol{x};\boldsymbol{d}).$$

类似于定理 4.3.2 的证明可得到结论. □

定理 4.3.4 设 f_1 和 f_2 同定理 4.3.2 的条件. 若 f_1 和 f_2 在 $\boldsymbol{x}\in\mathscr{B}$ 处是正则的，则

$$\partial^0(f_1+f_2)(\boldsymbol{x})=\partial^0 f_1(\boldsymbol{x})+\partial^0 f_2(\boldsymbol{x}).$$

证明 由定理 4.2.12 的(3)，有

$$(f_1+f_2)^0(\boldsymbol{x};\boldsymbol{d})=f_1^0(\boldsymbol{x};\boldsymbol{d})+f_2^0(\boldsymbol{x};\boldsymbol{d}),$$

再由定理 4.2.11 的(1)即可得证. □

下面是中值定理.

定理 4.3.5 设 $\boldsymbol{x},\boldsymbol{y}\in\mathscr{B}$，$S\subset\mathscr{B}$ 是开集，$[\boldsymbol{x},\boldsymbol{y}]\subset S$. 若 f 在 S 上是 Lipschitz 的，则存在点 $\boldsymbol{\xi}\in(\boldsymbol{x},\boldsymbol{y})$，使得

$$f(\boldsymbol{y})-f(\boldsymbol{x})\in\langle\partial^0 f(\boldsymbol{\xi}),\boldsymbol{y}-\boldsymbol{x}\rangle.$$

证明 设 $\lambda\in[0,1]$，记 $\boldsymbol{x}_\lambda=\boldsymbol{x}+\lambda(\boldsymbol{y}-\boldsymbol{x})$. 作函数 g：$[0,1]\to R$，$g(\lambda)=f(\boldsymbol{x}_\lambda)$，显然 g 在 $(0,1)$ 上是 Lipschitz 的. 令 $\delta=\pm 1$，我们有

$$\begin{aligned}
g^0(\lambda;\delta)&=\limsup_{\substack{\lambda'\to\lambda\\ t\to 0^+}}\frac{g(\lambda'+t\delta)-g(\lambda')}{t}\\
&=\limsup_{\substack{\lambda'\to\lambda\\ t\to 0^+}}\frac{f\big(\boldsymbol{x}+(\lambda'+t\delta)(\boldsymbol{y}-\boldsymbol{x})\big)-f\big(\boldsymbol{x}+\lambda'(\boldsymbol{y}-\boldsymbol{x})\big)}{t}\\
&\leqslant\limsup_{\substack{\boldsymbol{y}'\to\boldsymbol{x}_\lambda\\ t\to 0^+}}\frac{f\big(\boldsymbol{y}'+t\delta(\boldsymbol{y}-\boldsymbol{x})\big)-f(\boldsymbol{y}')}{t}\\
&=f^0\big(\boldsymbol{x}_\lambda;\delta(\boldsymbol{y}-\boldsymbol{x})\big)\\
&=\max\langle\partial^0 f(\boldsymbol{x}_\lambda),\delta(\boldsymbol{y}-\boldsymbol{x})\rangle.
\end{aligned}$$

于是，由定理 4.2.11 的(1)得 $\partial^0 g(\lambda) \subset \langle \partial^0 f(\boldsymbol{x}_\lambda), \boldsymbol{y}-\boldsymbol{x}\rangle$. 现考虑函数

$$\theta: [0, 1] \to R, \ \theta(\lambda) = g(\lambda) + \lambda\big[f(\boldsymbol{x}) - f(\boldsymbol{y})\big],$$

则有 $\theta(0) = \theta(1) = f(\boldsymbol{x})$. 因为 θ 在(0, 1)上是连续的，故它在(0, 1)上可达到局部极小值. 由此，据定理 4.2.3 得 $\boldsymbol{0} \in \partial^0\theta(\lambda)$，再由定理 4.3.1 和定理 4.3.2 有

$$\boldsymbol{0} \in \big\langle \partial^0 f(\boldsymbol{x}_\lambda), \boldsymbol{y}-\boldsymbol{x}\big\rangle + f(\boldsymbol{x}) - f(\boldsymbol{y}).$$

取 $\xi = \boldsymbol{x}_\lambda$，即得定理的结论. ▯

以下介绍复合函数求 G-次微分的链式法则.

定理 4.3.6　设 $\boldsymbol{h}: \mathscr{B} \to R^m$ 是向量函数，$\boldsymbol{h}(\boldsymbol{x}) = (h_1(\boldsymbol{x}), \cdots, h_m(\boldsymbol{x}))^T$，$g: R^m \to R$ 和 $f: \mathscr{B} \to R$ 是实值函数，$f(\boldsymbol{x}) = g(\boldsymbol{h}(\boldsymbol{x}))$. 若 $h_i(i = 1, \cdots, m)$ 在 $\boldsymbol{x} \in \mathscr{B}$ 处是 Lipschitz 的，g 在 $\boldsymbol{h}(\boldsymbol{x})$处是 Lipschitz 的，则

$$\partial^0 f(\boldsymbol{x}) \subset \overline{\mathrm{co}}\Big\{ \sum_{i=1}^m \lambda_i \boldsymbol{x}_i^* \,|\, \boldsymbol{x}_i^* \in \partial^0 h_i(\boldsymbol{x}), \ i = 1, \cdots, m;$$

$$\lambda \in \partial^0 g(\boldsymbol{h}(\boldsymbol{x})), \ \boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_m)^T \Big\}.$$

此外，使上式为等式，只要以下条件之一成立：

(1) g 在 $\boldsymbol{h}(\boldsymbol{x})$处是正则的，$h_i(i = 1, \cdots, m)$ 在 $\boldsymbol{x}$ 处是正则的，$\partial^0 g(\boldsymbol{h}(\boldsymbol{x}))$的元素的每一个分量是非负的(这时 f 在 $\boldsymbol{x}$ 处是正则的).

(2) g 在 $\boldsymbol{h}(\boldsymbol{x})$处是严格可微的.

(3) g 在 $\boldsymbol{h}(\boldsymbol{x})$处是正则的和 $\boldsymbol{h}$ 在 $\boldsymbol{x}$ 处是严格可微的(这时 f 在 $\boldsymbol{x}$ 处是正则的).

证明　记

$$S = \overline{\mathrm{co}}\Big\{ \sum_{i=1}^m \lambda_i \boldsymbol{x}_i^* \,|\, \boldsymbol{x}_i^* \in \partial^0 h_i(\boldsymbol{x}), \ i = 1, \cdots, m;$$

$$\lambda \in \partial^0 g\big(\boldsymbol{h}(\boldsymbol{x})\big),\ \boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_m)^T\Big\},$$

设

$$\sigma_S(\boldsymbol{d}) = \max\Big\{\sum_{i=1}^{m}\lambda_i\langle \boldsymbol{x}_i^*, \boldsymbol{d}\rangle \,|\, \boldsymbol{x}_i^* \in \partial^0 h_i(\boldsymbol{x}),$$

$$i = 1, \cdots, m;\ \boldsymbol{\lambda} \in \partial^0 g(\boldsymbol{h}(\boldsymbol{x}))\Big\}.$$

易知 σ_S 是 S 的支撑函数，由定理 4.2.11，只要证明 $\sigma_S(\boldsymbol{d}) \geqslant f^0(\boldsymbol{x}; \boldsymbol{d})$，就有 $\partial^0 f(\boldsymbol{x}) \subset S$. 设 $\varepsilon > 0$，作函数 σ_ε: $\mathscr{B} \to R$,

$$\sigma_\varepsilon(\boldsymbol{d}) = \max\Big\{\sum_{i=1}^{m}\lambda_i\langle \boldsymbol{x}_i^*, \boldsymbol{d}\rangle \,|\, \boldsymbol{x}_i^* \in \partial^0 h_i(\boldsymbol{x}_i),\ i = 1, \cdots, m;$$

$$\lambda \in \partial^0 g(\boldsymbol{\xi}),\ \boldsymbol{x}_i \in \boldsymbol{x} + \varepsilon B,\ \boldsymbol{\xi} \in \boldsymbol{h}(\boldsymbol{x}) + \varepsilon B\Big\}.$$

由定义 4.1.2，对任给的 ε，可以找到 $\boldsymbol{x}$ 附近的点 $\boldsymbol{x}'$，0 附近的 $t > 0$，使得

$$f^0(\boldsymbol{x}; \boldsymbol{d}) \leqslant \frac{f(\boldsymbol{x}' + t\boldsymbol{d}) - f(\boldsymbol{x}')}{t} + \varepsilon. \tag{4.3.2}$$

取 $\boldsymbol{x}'$, $\boldsymbol{x}' + t\boldsymbol{d} \in \boldsymbol{x} + \varepsilon B$，以及 $\boldsymbol{h}(\boldsymbol{x}')$, $\boldsymbol{h}(\boldsymbol{x}' + t\boldsymbol{d}) \in \boldsymbol{h}(\boldsymbol{x}) + \varepsilon B'$，其中 $B' \subset R^m$ 是开单位球. 由定理 4.3.5 得

$$f(\boldsymbol{x}' + t\boldsymbol{d}) - f(\boldsymbol{x}') = g\big(\boldsymbol{h}(\boldsymbol{x}' + t\boldsymbol{d})\big) - g\big(\boldsymbol{h}(\boldsymbol{x}')\big)$$

$$= \sum_{i=1}^{m}\lambda_i\big[h_i(\boldsymbol{x}' + t\boldsymbol{d}) - h_i(\boldsymbol{x}')\big],$$

其中 $(\lambda_1, \cdots, \lambda_m)^T \in \partial^0 g(\boldsymbol{\xi})$，$\boldsymbol{\xi}$ 是线段 $\big[\boldsymbol{h}(\boldsymbol{x}'), \boldsymbol{h}(\boldsymbol{x}' + t\boldsymbol{d})\big]$ 中的点，从而 $\boldsymbol{\xi} \in \boldsymbol{h}(\boldsymbol{x}) + \varepsilon B'$. 再利用定理 4.3.5 可得

$$f(\boldsymbol{x}' + t\boldsymbol{d}) - f(\boldsymbol{x}') = \sum_{i=1}^{m}\lambda_i\langle \boldsymbol{x}_i^*, t\boldsymbol{d}\rangle,$$

其中 $\boldsymbol{x}_i^* \in \partial^0 h_i(\boldsymbol{x}_i)$，$\boldsymbol{x}_i$ 是线段 $[\boldsymbol{x}', \boldsymbol{x}' + t\boldsymbol{d}]$ 中的点. 因而 $\boldsymbol{x}_i \in \boldsymbol{x} + \varepsilon B$. 将上式代入(4.3.2)，有

$$f^0(\boldsymbol{x};\ \boldsymbol{d}) \leqslant \sigma_\varepsilon(\boldsymbol{d}) + \varepsilon. \tag{4.3.3}$$

下面证明 $\varepsilon \to 0^+$ 时，$\sigma_\varepsilon(\boldsymbol{d}) \to \sigma_S(\boldsymbol{d})$. 设任给 $\delta > 0$, γ 是 $h_i(i = 1, \cdots, m)$ 的 Lipschitz 常数. 不难选择 ε，使得对每一 $h_i(i = 1, \cdots, m)$，它在 $\boldsymbol{x} + \varepsilon B$ 上是 γ 秩 Lipschitz 的. 对每一 $\boldsymbol{x}_i \in \boldsymbol{x} + \varepsilon B$，我们有

$$h_i^0(\boldsymbol{x}_i;\ \pm \boldsymbol{d}) \leqslant h_i^0(\boldsymbol{x};\ \pm \boldsymbol{d}) + \frac{\delta}{\gamma},\ i = 1, \cdots, m.$$

将上式两端同乘一个 $|\lambda_i|$，其中 $(\lambda_1, \cdots, \lambda_m)^T \in \partial^0 g(\xi)$，注意 $|\lambda_i| \leqslant \gamma$，则有

$$h_i^0(\boldsymbol{x}_i;\ \lambda_i\boldsymbol{d}) \leqslant h_i^0(\boldsymbol{x};\ \lambda_i\boldsymbol{d}) + \delta,\ i = 1, \cdots, m. \tag{4.3.4}$$

根据定理 4.2.2 的(4)，也可选择 ε 使得

$$\partial^0 g\big(\boldsymbol{h}(\boldsymbol{x})\big) + \varepsilon B \subset \partial^0 g\big(\boldsymbol{h}(\boldsymbol{x})\big) + \delta B'.$$

利用(4.3.4)和上式，有

$$\begin{aligned}
\sigma_\varepsilon(\boldsymbol{d}) &\leqslant \max\Big\{ \sum_{i=1}^m \max\big[\lambda_i\langle \boldsymbol{x}_i^*,\ \boldsymbol{d}\rangle \,|\, \boldsymbol{x}_i^* \in \partial^0 h_i(\boldsymbol{x}_i),\ \boldsymbol{x}_i \in \boldsymbol{x} + \varepsilon B\big] \\
&\qquad |\, (\lambda_1, \cdots, \lambda_m)^T \in \partial^0 g(\boldsymbol{h}(\boldsymbol{x})) + \varepsilon B' \Big\} \\
&\leqslant \max\Big\{ \sum_{i=1}^m \big(h_i^0(\boldsymbol{x},\ \lambda_i\boldsymbol{d}) + \delta\big) \\
&\qquad |\, (\lambda_1, \cdots, \lambda_m)^T \in \partial^0 g(\boldsymbol{h}(\boldsymbol{x})) + \delta B' \Big\} \\
&\leqslant \max\Big\{ \sum_{i=1}^m \max\big[\lambda_i\langle \boldsymbol{x}_i^*,\ \boldsymbol{d}\rangle \,|\, \boldsymbol{x}_i^* \in \partial^0 h_i(\boldsymbol{x})\big] \\
&\qquad |\, (\lambda_1, \cdots, \lambda_m)^T \in \partial^0 g(\boldsymbol{h}(\boldsymbol{x})) + \varepsilon B' \Big\} + m\delta \\
&\leqslant \sigma_S(\boldsymbol{d}) + n\delta\gamma \,\|\, \boldsymbol{d} \,\|\, + m\delta.
\end{aligned}$$

由 δ 的任意性和 $\sigma_S(\boldsymbol{d}) \leqslant \sigma_\varepsilon(\boldsymbol{d})$，可得 $\lim\limits_{\varepsilon \to 0^+} \sigma_\varepsilon(\boldsymbol{d}) = \sigma_S(\boldsymbol{d})$. 根据

(4.3.3)得到 $f^0(\boldsymbol{x};\boldsymbol{d})\leqslant\sigma_S(\boldsymbol{d})$，因此 $\partial^0 f(\boldsymbol{x})\subset S$.

以下证明等式成立的情形.

(1) 因为已知 $\partial^0 g(\boldsymbol{h}(\boldsymbol{x}))$的元素 $\boldsymbol{\lambda}=(\lambda_1,\cdots,\lambda_m)^T$ 的每一分量 $\lambda_i(i=1,\cdots,m)$ 为非负，我们有

$$\sigma_S(\boldsymbol{d})=\max\left\{\sum_{i=1}^{m}\max\left[\lambda_i\langle\boldsymbol{x}_i^*,\boldsymbol{d}\rangle\,|\,\boldsymbol{x}_i^*\in\partial^0 h_i(\boldsymbol{x})\right]\right.$$

$$\left.|\,(\lambda_1,\cdots,\lambda_m)^T\in\partial^0 g(\boldsymbol{h}(\boldsymbol{x}))\right\}$$

$$=\max\left\{\sum_{i=1}^{m}\lambda_i w_i\,|\,\boldsymbol{\lambda}\in\partial^0 g(\boldsymbol{h}(\boldsymbol{x}))\right\},$$

其中 $w_i=\max\left[\langle\boldsymbol{x}_i^*,\boldsymbol{d}\rangle\,|\,\boldsymbol{x}_i^*\in\partial^0 h_i(\boldsymbol{x})\right]\ (i=1,\cdots,m)$.

令 $\boldsymbol{w}=(w_1,\cdots,w_m)^T$. 从上式有

$$\sigma_S(\boldsymbol{d})=g'\big(\boldsymbol{h}(\boldsymbol{x});\boldsymbol{w}\big)=\lim_{t\to0^+}\frac{g(\boldsymbol{h}(\boldsymbol{x})+t\boldsymbol{w})-g\big(\boldsymbol{h}(\boldsymbol{x})\big)}{t},$$

或

$$\sigma_S(\boldsymbol{d})=\lim_{t\to0^+}\frac{g\big(\boldsymbol{h}(\boldsymbol{x})+t\boldsymbol{d}\big)-g\big(\boldsymbol{h}(\boldsymbol{x})\big)}{t}$$

$$-\frac{g\big(\boldsymbol{h}(\boldsymbol{x})+t\boldsymbol{w}\big)-g\big(\boldsymbol{h}(\boldsymbol{x}+t\boldsymbol{d})\big)}{t}.\qquad(4.3.5)$$

注意 g 在 $\boldsymbol{h}(\boldsymbol{x})$处是 Lipschitz 的，故有

$$\left|\frac{g\big(\boldsymbol{h}(\boldsymbol{x})+t\boldsymbol{w}\big)-g\big(\boldsymbol{h}(\boldsymbol{x}+t\boldsymbol{d})\big)}{t}\right|$$

$$\leqslant\gamma\left\|\boldsymbol{w}-\frac{\boldsymbol{h}(\boldsymbol{x}+t\boldsymbol{d})-\boldsymbol{h}(\boldsymbol{x})}{t}\right\|.$$

当 $t\to0^+$ 时，上式右端趋于 0. 因此，从(4.3.5)注意到 $f(\boldsymbol{x})=g\big(\boldsymbol{h}(\boldsymbol{x})\big)$，有 $\sigma_S(\boldsymbol{d})=f'(\boldsymbol{x};\boldsymbol{d})$. 前面已经证明 $f^0(\boldsymbol{x};\boldsymbol{d})\leqslant\sigma_S(\boldsymbol{d})$，由于$f'(\boldsymbol{x};\boldsymbol{d})\leqslant f^0(\boldsymbol{x};\boldsymbol{d})$，故得 $f^0(\boldsymbol{x};\boldsymbol{d})=f'(\boldsymbol{x};\boldsymbol{d})$. 由

此，$f^0(\boldsymbol{x};\boldsymbol{d})$是 S 的支撑函数，利用定理 4.2.11，便得到 $\partial^0 f(\boldsymbol{x}) = S$.

(2) 由已设 g 在 $\boldsymbol{h}(\boldsymbol{x})$ 处是严格可微的，不妨设 $\beta = D_S g\{\boldsymbol{h}(\boldsymbol{x})\} \geqslant 0$，则有

$$\sigma_S(\boldsymbol{d}) = \beta \boldsymbol{h}^0(\boldsymbol{x};\boldsymbol{d}) = \lim_{\substack{\boldsymbol{x}'\to\boldsymbol{x}\\ t\to 0^+}} \frac{\beta\{\boldsymbol{h}(\boldsymbol{x}'+t\boldsymbol{d}) - \boldsymbol{h}(\boldsymbol{x}')\}}{t}. \tag{4.3.6}$$

另外，由定义 4.2.3 得

$$\beta = \lim_{\substack{\boldsymbol{h}(\boldsymbol{x}'+t\boldsymbol{d})\to\boldsymbol{h}(\boldsymbol{x}')\\ \boldsymbol{x}'\to\boldsymbol{x}\\ t\to 0^+}} \frac{g\{\boldsymbol{h}(\boldsymbol{x}'+t\boldsymbol{d})\} - g\{\boldsymbol{h}(\boldsymbol{x}')\}}{\boldsymbol{h}(\boldsymbol{x}'+t\boldsymbol{d}) - \boldsymbol{h}(\boldsymbol{x}')}.$$

从上式和(4.3.6)有

$$\sigma_S(\boldsymbol{d}) = \lim_{\substack{\boldsymbol{x}'\to\boldsymbol{x}\\ t\to 0^+}} \frac{g\{\boldsymbol{h}(\boldsymbol{x}'+t\boldsymbol{d})\} - g\{\boldsymbol{h}(\boldsymbol{x}')\}}{t} = f^0(\boldsymbol{x};\boldsymbol{d}),$$

再根据定理 4.2.11，即得 $\partial^0 f(\boldsymbol{x}) = S$.

(3)与(1)的证明类似. □

定理 4.3.7　设 $\mathscr{D}$ 是 Banach 空间，φ: $\mathscr{B}\to\mathscr{D}$ 是映射，g: $\mathscr{D}\to R$ 是实值函数，$f = g\circ\varphi$. 若 φ 在 $\boldsymbol{x}\in B$ 处是严格可微的，g 在 $\varphi(\boldsymbol{x})$处是 Lipschitz 的，则 f 在 $\boldsymbol{x}$ 处是 Lipschitz 的，并且有

$$\partial^0 f(\boldsymbol{x}) \subset \partial^0 g\{\varphi(\boldsymbol{x})\}\circ D_S\varphi(\boldsymbol{x}). \tag{4.3.7}$$

再若 g 在 $\varphi(\boldsymbol{x})$处是正则的，则 f 在 $\boldsymbol{x}$ 处是正则的，并且(4.3.7)成为等式. 注意(4.3.7)等价于：若 $\boldsymbol{x}^*\in\partial^0 f(\boldsymbol{x})$，则存在 $\boldsymbol{y}^*\in\partial^0 g\{\varphi(\boldsymbol{x})\}$ 和 $D_S\varphi(\boldsymbol{x})$，使得对任意的 $\boldsymbol{d}\in\mathscr{B}$ 有

$$\langle\boldsymbol{x}^*,\boldsymbol{d}\rangle = \langle\boldsymbol{y}^*,\langle D_S\varphi(\boldsymbol{x}),\boldsymbol{d}\rangle\rangle.$$

证明　根据定理 4.2.4，φ 在 $\boldsymbol{x}$ 处是严格可微的，则 φ 在 $\boldsymbol{x}$ 处

是 Lipschitz 的. 于是,存在 $\delta_1 > 0$, 使得对任意的 $\boldsymbol{x}^1$, $\boldsymbol{x}^2 \in \boldsymbol{x} + \delta_1 B_{\mathscr{X}}$ 和某 $\gamma_1 > 0$ 有

$$\left\|\varphi(\boldsymbol{x}^1) - \varphi(\boldsymbol{x}^2)\right\|_{\mathscr{D}} \leqslant \gamma_1 \|\boldsymbol{x}^1 - \boldsymbol{x}^2\|_{\mathscr{X}}. \tag{4.3.8}$$

由于 g 在 $\varphi(\boldsymbol{x})$ 处是 Lipschitz 的,则存在 $\delta_2 > 0$, $\gamma_2 > 0$ 使得对任意的 $\boldsymbol{u}^1$, $\boldsymbol{u}^2 \in \varphi(\boldsymbol{x}) + \delta_2 B_{\mathscr{D}}$, 有

$$\left|g(\boldsymbol{u}^1) - g(\boldsymbol{u}^2)\right| \leqslant \gamma_2 \|\boldsymbol{u}^1 - \boldsymbol{u}^2\|_{\mathscr{D}}. \tag{4.3.9}$$

因为 φ 在 $\boldsymbol{x}$ 处是连续的,选择 $\delta > 0$, 使得当 $\delta \leqslant \delta_1$ 时,对任意的 $\boldsymbol{x}^1$, $\boldsymbol{x}^2 \in \boldsymbol{x} + \delta B_{\mathscr{X}}$ 有 $\varphi(\boldsymbol{x}^1)$, $\varphi(\boldsymbol{x}^2) \in \varphi(\boldsymbol{x}) + \delta_2 B_{\mathscr{D}}$. 由(4.3.8)和(4.3.9)我们有

$$\left|f(\boldsymbol{x}^1) - f(\boldsymbol{x}^2)\right| \leqslant \gamma_1\gamma_2 \|\boldsymbol{x}^1 - \boldsymbol{x}^2\|_{\mathscr{X}},$$

故 f 在 $\boldsymbol{x}$ 处是 Lipschitz 的.

设 $S = \partial^0 g(\varphi(\boldsymbol{x})) \circ D_S\varphi(\boldsymbol{x})$, 记 $\varphi' = D_S\varphi(\boldsymbol{x})$, 作

$$g^0\left(\varphi(\boldsymbol{x}); \varphi'(\boldsymbol{d})\right) = \max\left\{\langle \boldsymbol{y}^*, \varphi'(\boldsymbol{d})\rangle \mid \boldsymbol{y}^* \in \partial^0 g(\varphi(\boldsymbol{x}))\right\}.$$

因为 S 是凸的弱紧集, $g^0(\varphi(\boldsymbol{x}); \varphi'(\boldsymbol{d}))$ 是 S 的支撑函数,故有

$$\begin{aligned}\sigma_S(\boldsymbol{d}) &= \max\left\{\langle \boldsymbol{x}^*, \boldsymbol{d}\rangle \mid \boldsymbol{x}^* \in S\right\} \\ &= g^0\left(\varphi(\boldsymbol{x}); \varphi'(\boldsymbol{d})\right).\end{aligned}$$

现在只要证明 $f^0(\boldsymbol{x}; \boldsymbol{d}) \leqslant g^0(\varphi(\boldsymbol{x}); \varphi'(\boldsymbol{d}))$, 根据定理 4.2.11 即可得 $\partial^0 f(\boldsymbol{x}) \subset S$. 为此对任给的 $\varepsilon > 0$, 选择

$$\boldsymbol{x}', \ \boldsymbol{x}' + t\boldsymbol{d} \in \boldsymbol{x} + \varepsilon B_{\mathscr{X}},$$

$$\varphi(\boldsymbol{x}'), \ \varphi(\boldsymbol{x}' + t\boldsymbol{d}) \in \varphi(\boldsymbol{x}) + \varepsilon B_{\mathscr{D}},$$

由 $f^0(\boldsymbol{x}; \boldsymbol{d})$ 的定义,使得

$$f^0(\boldsymbol{x}; \boldsymbol{d}) \leqslant \frac{f(\boldsymbol{x}' + t\boldsymbol{d}) - f(\boldsymbol{x}')}{t} + \varepsilon. \tag{4.3.10}$$

另外，由定理 4.3.5 有

$$f(\boldsymbol{x}'+t\boldsymbol{d})-f(\boldsymbol{x}')=g\big(\varphi(\boldsymbol{x}'+t\boldsymbol{d})\big)-g\big(\varphi(\boldsymbol{x}')\big)$$
$$=\langle \boldsymbol{y}^*,\ \varphi(\boldsymbol{x}'+t\boldsymbol{d})-\varphi(\boldsymbol{x}')\rangle,$$

其中 $\boldsymbol{y}^*\in\partial^0 g(\xi)$，$\xi\in\big[\varphi(\boldsymbol{x}'),\ \varphi(\boldsymbol{x}'+t\boldsymbol{d})\big]$，因此 $\xi\in\varphi(\boldsymbol{x})+\varepsilon B_{\mathscr{Y}}$. 令 $\boldsymbol{w}=\varphi(\boldsymbol{x}'+t\boldsymbol{d})-\varphi(\boldsymbol{x}')$，由定理 4.1.2 知 $g^0\big(\varphi(\boldsymbol{x});\ \boldsymbol{w}\big)$ 在 $\big(\varphi(\boldsymbol{x}),\ \boldsymbol{w}\big)$ 处是上半连续的. 选择适当的 ε，使得 $\xi\in\varphi(\boldsymbol{x})+\varepsilon B_{\mathscr{Y}}$，有

$$g^0(\xi;\ \boldsymbol{w})\leqslant g^0\big(\varphi(\boldsymbol{x});\ \boldsymbol{w}\big)+\varepsilon t.$$

由(4.3.10)知

$$f^0(\boldsymbol{x};\ \boldsymbol{d})\leqslant\left\langle \boldsymbol{y}^*,\ \frac{\boldsymbol{w}}{t}\right\rangle+\varepsilon$$
$$\leqslant\max\left\{\left\langle \boldsymbol{y}^*,\ \frac{\boldsymbol{w}}{t}\right\rangle\middle|\ \boldsymbol{y}^*\in\partial^0 g\big(\varphi(\boldsymbol{x})\big)\right\}+\varepsilon,$$

或

$$f^0(\boldsymbol{x};\ \boldsymbol{d})\leqslant g^0\left(\xi;\ \frac{\boldsymbol{w}}{t}\right)+\varepsilon\leqslant g^0\left(\varphi(\boldsymbol{x});\ \frac{\boldsymbol{w}}{t}\right)+2\varepsilon.$$

上式即为

$$f^0(\boldsymbol{x};\ \boldsymbol{d})\leqslant\max\left\{\left\langle \boldsymbol{y}^*,\ \frac{\varphi(\boldsymbol{x}'+t\boldsymbol{d})-\varphi(\boldsymbol{x}')}{t}\right\rangle\middle|\ \boldsymbol{y}^*\in\partial^0 g\big(\varphi(\boldsymbol{x})\big)\right\}$$
$$+2\varepsilon.$$

令 $\boldsymbol{x}'\to\boldsymbol{x}$，$t\to0^+$，得

$$f^0(\boldsymbol{x};\ \boldsymbol{d})\leqslant\max\{\langle \boldsymbol{y}^*,\ \varphi'(\boldsymbol{d})\rangle\mid \boldsymbol{y}^*\in\partial^0 g(\varphi(\boldsymbol{x}))\}+2\varepsilon.$$

因为 ε 可任意小，故得

$$f^0(\boldsymbol{x};\ \boldsymbol{d})\leqslant g^0\big(\varphi(\boldsymbol{x});\ \varphi'(\boldsymbol{d})\big).\tag{4.3.11}$$

若 g 在 $\varphi(\boldsymbol{x})$ 处是正则的，则因为

$$\frac{[\varphi(\boldsymbol{x}+t\boldsymbol{d})-\varphi(\boldsymbol{x})-t\varphi'(\boldsymbol{d})]}{t}\to 0\ (t\to 0),$$

我们有

$$\lim_{t\to 0^+}\frac{g(\varphi(\boldsymbol{x})+t\varphi'(\boldsymbol{d}))-g(\varphi(\boldsymbol{x}))}{t}=\lim_{t\to 0^+}\frac{g(\varphi(\boldsymbol{x}+t\boldsymbol{d}))-g(\varphi(\boldsymbol{x}))}{t}$$

$$=\lim_{t\to 0^+}\frac{f(\boldsymbol{x}+t\boldsymbol{d})-f(\boldsymbol{x})}{t}$$

$$=f'(\boldsymbol{x};\boldsymbol{d})\leqslant f^0(\boldsymbol{x};\boldsymbol{d}).$$

从上式和(4.3.11)得到 $f^0(\boldsymbol{x};\boldsymbol{d})=g^0(\varphi(\boldsymbol{x});\varphi'(\boldsymbol{d}))$，故 $\partial^0 f(\boldsymbol{x})=S$. ▯

推论 4.3.8 设 $f_i:\mathscr{B}\to R\ (i=1,\cdots,m)$ 在 $\boldsymbol{x}\in\mathscr{B}$ 处是 Lipschitz 的，$f:\mathscr{B}\to R$，$f(\boldsymbol{x})=\max\{f_i(\boldsymbol{x})\,|\,i=1,\cdots,m\}$，

$$I(\boldsymbol{x})=\{i\in\{1,\cdots,m\}\,|\,f_i(\boldsymbol{x})=f(\boldsymbol{x})\},$$

则

$$\partial^0 f(\boldsymbol{x})\subset \mathrm{co}\{\partial^0 f_i(\boldsymbol{x})\,|\,i\in I(\boldsymbol{x})\}.\tag{4.3.12}$$

再若当 $i\in I(\boldsymbol{x})$ 时 f_i 在 $\boldsymbol{x}$ 处是正则的，则(4.3.12)成为等式，并且 f 在 $\boldsymbol{x}$ 处是正则的.

证明 定义函数 $g:R^m\to R$，

$$g(u_1,\cdots,u_m)=\max\{u_i\,|\,i=1,\cdots,m\},$$

以及 $\boldsymbol{h}:\mathscr{B}\to R^m$，$\boldsymbol{h}(\boldsymbol{x})=(f_1(\boldsymbol{x}),\cdots,f_m(\boldsymbol{x}))^T$，则 $f(\boldsymbol{x})=g(\boldsymbol{h}(\boldsymbol{x}))$. 据定理 4.3.6 和例 4.2.2 即得结论. ▯

下面阐述乘法规则和除法规则.

定理 4.3.9 设 $f_1,f_2:\mathscr{B}\to R$ 在 $\boldsymbol{x}\in\mathscr{B}$ 处是 Lipschitz 的，则 $f_1\cdot f_2$ 在 $\boldsymbol{x}$ 处是 Lipschitz 的，并且

$$\partial^0(f_1\cdot f_2)(\boldsymbol{x})\subset f_2(\boldsymbol{x})\partial^0 f_1(\boldsymbol{x})+f_1(\boldsymbol{x})\partial^0 f_2(\boldsymbol{x}).\tag{4.3.13}$$

再若 $f_1(\boldsymbol{x}) \geqslant 0$, $f_2(\boldsymbol{x}) \geqslant 0$，并且 f_1 和 f_2 均在 $\boldsymbol{x}$ 处是正则的，则 $f_1 \cdot f_2$ 在 $\boldsymbol{x}$ 处是正则的，并且(4.3.13)成为等式.

证明　作函数 $g: R^2 \to R$, $g(u_1, u_2) = u_1 u_2$，以及 $\boldsymbol{h}: \mathscr{B} \to R^2$, $\boldsymbol{h}(\boldsymbol{x}) = \big(f_1(\boldsymbol{x}), f_2(\boldsymbol{x})\big)^T$，则 $(f_1 \cdot f_2)(\boldsymbol{x}) = g\big(\boldsymbol{h}(\boldsymbol{x})\big)$. 根据定理 4.3.6 可以推得结论. □

定理 4.3.10　设 $f_1, f_2: \mathscr{B} \to R$ 在 $\boldsymbol{x} \in \mathscr{B}$ 处是 Lipschitz 的，$f_2(\boldsymbol{x}) \neq 0$，则 f_1/f_2 在 $\boldsymbol{x}$ 处是 Lipschitz 的，并且有

$$\partial^0 (f_1/f_2)(\boldsymbol{x}) \subset \frac{f_2(\boldsymbol{x})\partial^0 f_1(\boldsymbol{x}) - f_1(\boldsymbol{x})\partial^0 f_2(\boldsymbol{x})}{f_2^2(\boldsymbol{x})}. \tag{4.3.14}$$

再若 $f_1(\boldsymbol{x}) \geqslant 0$, $f_2(\boldsymbol{x}) \geqslant 0$，并且 f_1 和 $-f_2$ 均在 $\boldsymbol{x}$ 处是正则的，则 f_1/f_2 在 $\boldsymbol{x}$ 处是正则的，并且(4.3.14)成为等式.

证明　由定理 4.3.6 可以推得定理成立. □

最后，我们介绍 G-偏次微分的概念，并且给出两个结论.

设 $\mathscr{B}_1$ 和 $\mathscr{B}_2$ 是 Banach 空间，$\mathscr{B} = \mathscr{B}_1 \times \mathscr{B}_2$, $f: \mathscr{B} \to R$, $\boldsymbol{x}_1 \in \mathscr{B}_1$, $\boldsymbol{x}_2 \in \mathscr{B}_2$, $f(\boldsymbol{x}_1, \boldsymbol{x}_2)$在$(\boldsymbol{x}_1, \boldsymbol{x}_2)$处是 Lipschitz 的. 固定 $\boldsymbol{x}_2$，记 $f_1^0(\boldsymbol{x}_1, \boldsymbol{x}_2; \boldsymbol{d}^1)$表示 f 在 $\boldsymbol{x}_1$ 处沿方向 $\boldsymbol{d}^1 \in \mathscr{B}_1$ 的 G-方向导数，则称$f(\cdot, \boldsymbol{x}_2)$ 在 $\boldsymbol{x}_1$ 处的 G-次微分 $\partial_1^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2)$为 f 在 $\boldsymbol{x}_1$ 处的 G-偏次微分. 固定 $\boldsymbol{x}_1$, $f_2^0(\boldsymbol{x}_1, \boldsymbol{x}_2; \boldsymbol{d}^2)$表示 f 在 $\boldsymbol{x}_2$ 处沿方向 $\boldsymbol{d}^2 \in \mathscr{B}_2$ 的 G-方向导数，则称$f(x_1, \cdot)$在 $\boldsymbol{x}_2$ 处的 G-次微分 $\partial_2^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2)$为 f 在 $\boldsymbol{x}_2$ 处的 G-偏次微分.

定理 4.3.11　设 f 在 $(\boldsymbol{x}_1, \boldsymbol{x}_2) \in \mathscr{B}_1 \times \mathscr{B}_2$ 处是正则的，则

$$\partial^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2) \subset \partial_1^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2) \times \partial_2^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2).$$

证明　设 $\boldsymbol{x}^* = (\boldsymbol{x}_1^*, \boldsymbol{x}_2^*) \in \partial^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2)$，由定理 4.2.1 的(2)有

$$f^0\big(\boldsymbol{x}_1, \boldsymbol{x}_2; (\boldsymbol{d}, \boldsymbol{0})\big)$$

$$= \max\big\{\langle(\boldsymbol{x}_1^*, \boldsymbol{x}_2^*), (\boldsymbol{d}, \boldsymbol{0})\rangle \,\big|\, \boldsymbol{x}^* \in \partial^0 f(\boldsymbol{x})$$

$\geqslant \langle \boldsymbol{x}_1^*, \boldsymbol{d} \rangle\ \forall \boldsymbol{d} \in \mathscr{B} \}$.

因为 f 在 $\boldsymbol{x}$ 处是正则的，故有

$$f_1'(\boldsymbol{x}_1, \boldsymbol{x}_2; \boldsymbol{d}) = f'\big(\boldsymbol{x}_1, \boldsymbol{x}_2; (\boldsymbol{d}, \boldsymbol{0})\big)$$

$$= f^0(\boldsymbol{x}_1, \boldsymbol{x}_2; (\boldsymbol{d}, \boldsymbol{0})),$$

以及

$$f_1'(\boldsymbol{x}_1, \boldsymbol{x}_2; \boldsymbol{d}) \geqslant \langle \boldsymbol{x}_1^*, \boldsymbol{d} \rangle \quad \forall \boldsymbol{d} \in \mathscr{B}.$$

由此，我们有 $\boldsymbol{x}_1^* \in \partial_1^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2)$. 同理，可得 $\boldsymbol{x}_2^* \in \partial_2^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2)$. ▯

定理 4.3.12 设 $\mathscr{B}_1^*$ 和 $\mathscr{B}_2^*$ 分别是 $\mathscr{B}_1$ 和 $\mathscr{B}_2$ 的对偶空间. 记

$$\pi_1 \partial^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2)$$

$$= \{\boldsymbol{x}_1^* \in \mathscr{B}_1^* \mid (\boldsymbol{x}_1^*, \boldsymbol{x}_2^*) \in \partial^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2),\ \boldsymbol{x}_2^* \in \mathscr{B}_2^* \},$$

$$\pi_2 \partial^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2)$$

$$= \{\boldsymbol{x}_2^* \in \mathscr{B}_2^* \mid (\boldsymbol{x}_1^*, \boldsymbol{x}_2^*) \in \partial^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2),\ \boldsymbol{x}_1^* \in \mathscr{B}_1^* \},$$

则

$$\partial_1^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2) \subset \pi_1 \partial^0 (\boldsymbol{x}_1, \boldsymbol{x}_2),$$

$$\partial_2^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2) \subset \pi_2 \partial^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2).$$

证明 固定 $\boldsymbol{x}_2$，作函数 $f_1: \mathscr{B}_1 \to R$, $f_1(\boldsymbol{x}) = f(\boldsymbol{x}_1, \boldsymbol{x}_2)$，以及映射 $F: \mathscr{B}_1 \to \mathscr{B}_1 \times \mathscr{B}_2$, $F(\boldsymbol{x}) = (\boldsymbol{x}_1, \boldsymbol{x}_2)$，则 $\langle D_S F(\boldsymbol{x}), \boldsymbol{d} \rangle = (\boldsymbol{d}, \boldsymbol{0})$. 因为 $f_1 = f \circ F$，利用定理 4.3.7 即得

$$\partial_1^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2) \subset \pi_1 \partial^0 f(\boldsymbol{x}_1, \boldsymbol{x}_2).$$

同理，可证第 2 个结论成立. ▯

例 4.3.1 已知平面上的 $N+1$ 个点：$(0, 0)$, $(1, 1)$, $(2, 2)$, …, $(N-1, N-1)$ 和 $(N, 0)$，考虑如何寻找一条直线来拟合这些点(图 4.3.1). 设直线方程为 $y = \alpha x + \beta$，问题归结为求以下的极小化问题：

$$\min f(\alpha,\beta)=\sum_{i=0}^{N-1}|\alpha i+\beta-i|+|\alpha N+\beta-0|.$$

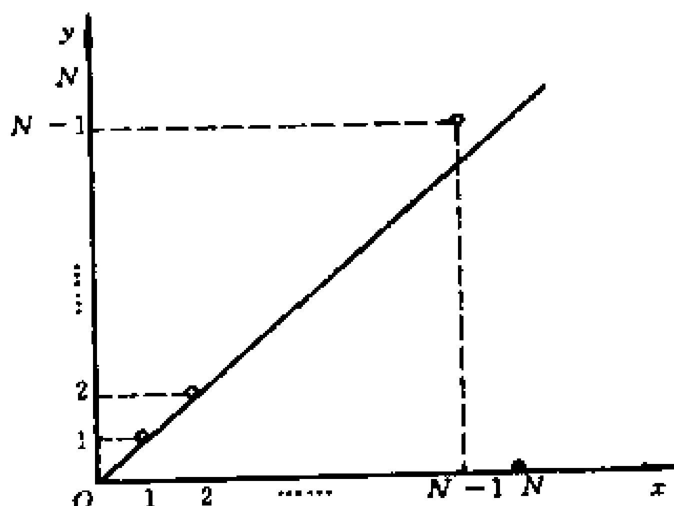


图　4.3.1

作函数 $f_{a,c}: R^2\to R$, $f_{a,c}(\alpha,\beta)=|a\alpha+\beta-c|$，它是以下两函数的复合：$g:R\to R$，　$h:R^2\to R$，

$$g(u)=|u|,\ h(\alpha,\beta)=a\alpha+\beta-c.$$

注意，h 是严格可微的，g 是凸函数，并且是正则的，由

$$\partial^0 g(u)=\begin{cases}1, & u>0;\\ [-1,1], & u=0;\\ -1, & u<0,\end{cases}$$

以及 $D_s h(\alpha,\beta)=(a,1)^T$，利用定理 4.3.7 得

$$\partial^0 f_{a,c}(\alpha,\beta)=\begin{cases}(a,1)^T, & a\alpha+\beta-c>0;\\ \{\gamma(a,1)^T \mid |\gamma|\leqslant 1\}, & a\alpha+\beta-c=0;\\ (-a,-1)^T, & a\alpha+\beta-c<0.\end{cases}\tag{4.3.15}$$

由于 $f(\alpha,\beta)=\sum_{i=0}^{N-1}f_{i,i}(\alpha,\beta)+f_{N,0}(\alpha,\beta)$，根据定理 4.3.4 有

$$\partial^0 f(\alpha, \beta) = \sum_{i=0}^{N-1} \gamma_i (i, 1)^T + \gamma_N (N, 1)^T,$$

其中 $|\gamma_i| \leqslant 1$ $(i = 0, 1, \cdots, N-1)$. 对于直线 $y = x$ (即 $\alpha = 1$, $\beta = 0$), 则当 $a = N$, $c = 0$ 时,有 $\alpha a + \beta - c = N > 0$, 由 (4.3.15)知 $\gamma_N = 1$. 从而

$$\partial^0 f(1, 0) = \sum_{i=0}^{N-1} \gamma_i (i, 1)^T + \gamma_N (N, 1)^T. \qquad (4.3.16)$$

由定理 4.2.3, (1, 0)是 $f(\alpha, \beta)$的极小点的必要条件是 $(0, 0)^T \in \partial^0 f(1, 0)$, 因为 f 是凸的,根据定理 3.4.1,它也是充分条件. 据此,由(4.3.16)得知 $y = x$ 是问题的解的充要条件为

$$(0, 0)^T \in \sum_{i=0}^{N-1} \gamma_i (i, 1)^T + \gamma_N (N, 1)^T, \qquad (4.3.17)$$

其中 $|\gamma_i| \leqslant 1$ $(i = 0, 1, \cdots, N-1)$. 容易看出,当 $N \geqslant 3$ 时,存在 $\gamma_i (i = 0, 1, \cdots, N-1)$ 满足(4.3.17).

§4.4 几 何 特 性

这一节,利用集合的距离函数,引进集合在一点处的 G-切锥和 G-法锥的概念. 由此,来刻画 G-次微分的有关几何特性.

设 $\mathscr{B}$ 是 Banach 空间,集合 $S \subset \mathscr{B}$ 非空,考虑 S 上的距离函数

$$d_S(\boldsymbol{x}) = \inf_{y \in S} \| \boldsymbol{x} - \boldsymbol{y} \|, \ \boldsymbol{x} \in \mathscr{B}. \qquad (4.4.1)$$

若 S 是闭的,则 $\boldsymbol{x} \in S$ 当且仅当 $d_S(\boldsymbol{x}) = 0$.

定理 4.4.1 设集合 $S \subset \mathscr{B}$ 非空,则距离函数 $d_S: \mathscr{B} \to R$ 在 $\mathscr{B}$ 上是 Lipschitz 的,即有

$$\left| d_S(\boldsymbol{x}^1) - d_S(\boldsymbol{x}^2) \right| \leqslant \| \boldsymbol{x}^1 - \boldsymbol{x}^2 \| \quad \forall \boldsymbol{x}^1, \boldsymbol{x}^2 \in \mathscr{B}. \qquad (4.4.2)$$

证明　由(4.4.1),对于任给的 $\epsilon > 0$, 存在 $\bar{\boldsymbol{y}} \in S$ 使得

$$d_S(\boldsymbol{x}^2) \geqslant \| \boldsymbol{x}^2 - \bar{\boldsymbol{y}} \| - \epsilon.$$

于是,我们有

$$\begin{aligned} d_S(\boldsymbol{x}^1) &\leqslant \| \boldsymbol{x}^1 - \bar{\boldsymbol{y}} \| \leqslant \| \boldsymbol{x}^1 - \boldsymbol{x}^2 \| + \| \boldsymbol{x}^2 - \bar{\boldsymbol{y}} \| \\ &\leqslant \| \boldsymbol{x}^1 - \boldsymbol{x}^2 \| + d_S(\boldsymbol{x}^2) + \epsilon. \end{aligned}$$

由 ϵ 的任意性得(4.4.2)成立. ▯

设 $\mathscr{B}^*$ 是 $\mathscr{B}$ 的对偶空间.

定义 4.4.1　设集合 $S \subset \mathscr{B}$ 非空,点 $\boldsymbol{x}^0 \in \mathscr{B}$.

(1) 称集合

$$T_S^0(\boldsymbol{x}^0) = \{\boldsymbol{d} \in \mathscr{B} \mid d_S^0(\boldsymbol{x}^0; \boldsymbol{d}) = 0\} \qquad (4.4.3)$$

是集合 S 在点 $\boldsymbol{x}^0$ 处的广义切锥,或 G-切锥.

(2) 称集合

$$N_S^0(\boldsymbol{x}^0) = \{\boldsymbol{x}^* \in \mathscr{B}^* \mid \langle \boldsymbol{x}^*, \boldsymbol{d} \rangle \leqslant 0 \quad \forall \boldsymbol{d} \in T_S^0(\boldsymbol{x}^0)\} \qquad (4.4.4)$$

是集合 S 在点 $\boldsymbol{x}^0$ 处的广义法锥,或 G-法锥.

注 4.4.1　设 $\boldsymbol{x} \subset \mathscr{B}$, 由定义 4.4.1 易知 $T_S^0(\boldsymbol{x})$是闭凸集.

定理 4.4.2　设集合 $S \subset \mathscr{B}$ 非空, $\boldsymbol{x} \in \mathscr{B}$, 则

$$N_S^0(\boldsymbol{x}) = \mathrm{cl} \bigcup_{\lambda \geqslant 0} \lambda \partial^0 d_S(\boldsymbol{x}), \qquad (4.4.5)$$

其中 cl 表示弱闭.

证明　由定义 4.4.1 的(1)和定理 4.2.1 的(2)可知, $\boldsymbol{d} \in T_S^0(\boldsymbol{x})$ 当且仅当 $\langle \boldsymbol{x}^*, \boldsymbol{d} \rangle \leqslant 0$ 对任意的 $\boldsymbol{x}^* \in \partial^0 d_S(\boldsymbol{x})$ 成立. 因为 $T_S^0(\boldsymbol{x})$是弱闭凸锥,故(4.4.5)成立. ▯

定理 4.4.3　设函数 $f: \mathscr{B} \to R$ 在集合 $X \subset \mathscr{B}$ 上是 γ 秩 Lipschitz 的, $\tilde{\boldsymbol{x}} \in S \subset X$. 若 f 在 $\tilde{\boldsymbol{x}} \in S$ 处达到极小值,则对任意的 $\tilde{\gamma} \geqslant \gamma$, 函数 $F(\boldsymbol{y}) = f(\boldsymbol{y}) + \tilde{\gamma}\, d_S(\boldsymbol{y})$ 在 $\tilde{\boldsymbol{x}} \in S$ 处达到极小值. 若

$\tilde{\gamma}>\gamma$, S 是闭的，并且 $f(\tilde{y})=\min\limits_{x\in X}F(x)$，则 $\tilde{y}\in S$.

证明 用反证法. 假设存在 $\bar{y}\in X$ 和 $\bar{\varepsilon}>0$，使得

$$f(\bar{y})+\tilde{\gamma}\,d_S(\bar{y})<f(\bar{x})-\tilde{\gamma}\,\bar{\varepsilon}.$$

取 $x'\in S$ 使 $\|\bar{y}-x'\|\leqslant d_S(\bar{y})+\bar{\varepsilon}$，由于 f 是 γ 秩 Lipschitz 的，则有

$$\begin{aligned}f(x')&\leqslant f(\bar{y})+\gamma\|\bar{y}-x'\|\\&\leqslant f(\bar{y})+\tilde{\gamma}\,(d_S(\bar{y})+\bar{\varepsilon})<f(\bar{x}).\end{aligned}$$

上式与 f 在 $\bar{x}\subset S$ 处达到极小值相矛盾. 现在设 $\tilde{\gamma}>\gamma$，函数 $F(y)=f(y)+\tilde{\gamma}\,d_S(y)$ 在 X 上的点 $\tilde{y}$ 处达到极小值，由前面所证，可知有

$$f(\tilde{y})+\tilde{\gamma}\,d_S(\tilde{y})=f(\bar{x})\leqslant f(\bar{x})+\frac{\gamma+\tilde{\gamma}}{2}\delta_S(\tilde{y}).$$

从上式得知 $d_S(\tilde{y})=0$，故 $\tilde{y}\in S$. ▯

推论 4.4.4 设函数 $f:\mathscr{B}\to R$ 在集合 $S\subset\mathscr{B}$ 上是 Lipschitz 的. 若 $f(\bar{x})=\min\limits_{x\in S}f(x)$，则

$$\boldsymbol{0}\in\partial^0 f(\bar{x})+N_S^0(\bar{x}).$$

证明 设 $U(\bar{x})$ 是 $\bar{x}$ 的邻域，f 在 $U(\bar{x})$ 上是 γ 秩 Lipschitz 的. 不妨设 $S\subset U(\bar{x})$，则由定理 4.4.3 知 $f(y)+\gamma d_S(y)$ 在 $\bar{x}$ 处达到极小值. 因此，有

$$\boldsymbol{0}\in\partial^0(f+\gamma d_S)(\bar{x})\subset\partial^0 f(\bar{x})+\gamma\partial^0 d_S(\bar{x})$$

$$\subset\partial^0 f(\bar{x})+N_S^0(\bar{x}).\ ▯$$

定理 4.4.5 设 $S\subset\mathscr{B}$ 是非空凸集，$x\in\mathscr{B}$，$N_S(x)$ 是 S 在点 x 处的法锥，即

$$N_S(x)=\{x^*\in\mathscr{B}^*\,|\,\langle x^*,y-x\rangle\leqslant 0\quad\forall y\in S\},$$

则 $N_S^0(x)=N_S(x)$.

证明 设 $\boldsymbol{x}^* \in N_S(\boldsymbol{x})$，$f(\boldsymbol{y}) = \langle \boldsymbol{x}^*, \boldsymbol{x} - \boldsymbol{y} \rangle$，我们有 $f(\boldsymbol{x}) = \min_{\boldsymbol{y} \in S} f(\boldsymbol{y})$. 于是，根据推论 4.4.4 知 $\boldsymbol{0} \in -\boldsymbol{x}^* + N_S^0(\boldsymbol{x})$，故得 $\boldsymbol{x}^* \in N_S^0(\boldsymbol{x})$.

反之，设 $\boldsymbol{x}^* \in N_S^0(\boldsymbol{x})$，令 $\boldsymbol{x}, \boldsymbol{x}' \in \mathscr{B}$，$\lambda \in (0, 1)$. 对任给的 $\varepsilon > 0$，选择 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in S$ 使得

$$\| \boldsymbol{x}^1 - \boldsymbol{x} \| \leqslant d_S(\boldsymbol{x}) + \varepsilon, \quad \| \boldsymbol{x}^2 - \boldsymbol{x}' \| \leqslant d_S(\boldsymbol{x}') + \varepsilon,$$

则

$$\begin{aligned} d_S(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{x}') &\leqslant \| \lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 - \lambda \boldsymbol{x} - (1-\lambda)\boldsymbol{x}' \| \\ &\leqslant \lambda d_S(\boldsymbol{x}) + (1-\lambda) d_S(\boldsymbol{x}') + \varepsilon. \end{aligned}$$

因为 ε 是任意的，故 d_S 是凸函数，于是由定理 4.4.2 得 $\boldsymbol{x}^* \in \partial^0 d_S(\boldsymbol{x})$. 利用定理 4.2.10，知

$$\langle \boldsymbol{x}^*, \boldsymbol{y} - \boldsymbol{x} \rangle \leqslant d_S(\boldsymbol{y}) - d_S(\boldsymbol{x}) \quad \forall \boldsymbol{y} \in \mathscr{B}.$$

因为对任意的 $\boldsymbol{y} \in S$，有 $\langle \boldsymbol{x}^*, \boldsymbol{y} - \boldsymbol{x} \rangle \leqslant 0$，所以 $\boldsymbol{x}^* \in N_S(\boldsymbol{x})$. ▯

注 4.4.2 设 $S \subset \mathscr{B}$ 是非空凸集，$\boldsymbol{x} \in \mathscr{B}$. 由定义 3.2.2，定义 4.4.1 和定理 4.4.5 可知，也有 $T_S^0(\boldsymbol{x}) = T_S(\boldsymbol{x})$.

推论 4.4.6 设 $S \subset \mathscr{B}$ 是非空凸集，$\boldsymbol{x} \in \mathscr{B}$，则 $\boldsymbol{d} \in T_S^0(\boldsymbol{x})$ 当且仅当 $d_S^0(\boldsymbol{x}; \boldsymbol{d}) = d_S'(\boldsymbol{x}; \boldsymbol{d}) = 0$.

证明 因为 d_S 是凸函数，由定理 4.2.12 的(2)知它是正则的，从而有 $d_S^0(\boldsymbol{x}; \boldsymbol{d}) = d_S'(\boldsymbol{x}; \boldsymbol{d}) = 0$. ▯

定理 4.4.7 设集合 $S \subset \mathscr{B}$ 非空，$\boldsymbol{x} \in \mathscr{B}$，则 $\boldsymbol{d} \in T_S^0(\boldsymbol{x})$ 当且仅当对任意的点列 $\{\boldsymbol{x}^k\} \subset S$，$\boldsymbol{x}^k \to \boldsymbol{x}$，$\{t_k\} \subset (0, +\infty)$，$t_k \to 0$ $(k \to \infty)$，存在序列 $\{\boldsymbol{d}^k\} \subset \mathscr{B}$，$\boldsymbol{d}^k \to \boldsymbol{d}$，使得对任何 $k = 1, 2, \cdots$，有 $\boldsymbol{x}^k + t_k \boldsymbol{d}^k \in S$.

证明 设 $\boldsymbol{d} \in T_S^0(\boldsymbol{x})$，$\{\boldsymbol{x}^k\} \subset S$，$\boldsymbol{x}^k \to \boldsymbol{x}$，$\{t_k\} \subset (0, +\infty)$，$t_k \to 0 (k \to \infty)$. 因为 $d_S^0(\boldsymbol{x}; \boldsymbol{d}) = 0$，$d_S$ 是凸函数，故有

$$\lim_{k \to \infty} \frac{d_S(\boldsymbol{x}^k + t_k \boldsymbol{d}) - d_S(\boldsymbol{x}^k)}{t_k} = \lim_{k \to \infty} \frac{d_S(\boldsymbol{x}^k + t_k \boldsymbol{d})}{t_k} = 0. \tag{4.4.6}$$

选择点 $y^k \in S$，使得

$$\| x^k + t_k d - y^k \| \leqslant d_S(x^k + t_k d) + t_k/k. \qquad (4.4.7)$$

设 $d^k = \dfrac{y^k - x^k}{t_k}$，由(4.4.6)和(4.4.7)，令 $k \to \infty$，则有 $d^k \to d$. 故对任何 $k = 1, 2, \cdots$，有 $x^k + t_k d^k \in S$.

反之，设 $\{x^k\} \subset S$，$x_k \to x$，$\{t_k\} \subset (0, +\infty)$，$t_k \to 0 (k \to \infty)$，有

$$d_S^0(x; d) = \lim_{k \to \infty} \frac{d_S(x^k + t_k d) - d_S(x^k)}{t_k}. \qquad (4.4.8)$$

选择 $y^k \in S$，使得

$$\| y^k - x^k \| \leqslant d_S(x^k) + t_k/k. \qquad (4.4.9)$$

在上式中令 $k \to \infty$，有 $y^k \to x$，故存在序列 $\{d^k\} \subset \mathscr{B}$，$d^k \to d$，使得 $y^k + t_k d^k \in S$. 据定理 4.4.1，我们有

$$d_S(x^k + t_k d) \leqslant d_S(y^k + t_k d^k) + \| x^k - y^k \| + t_k \| d^k - d \|.$$

由(4.4.9)得

$$d_S(x^k + t_k d) \leqslant d_S(x^k) + t_k(\| d^k - d \| + 1/k).$$

利用上式和(4.4.8)，令 $k \to \infty$，有 $d_S^0(x; d) = 0$，故 $d \in T_S^0(x)$. □

推论 4.4.8 设 $\mathscr{B}_1$，$\mathscr{B}_2$ 是 Banach 空间，$S_1 \subset \mathscr{B}_1$，$S_2 \subset \mathscr{B}_2$，$x = (x_1, x_2) \in S_1 \times S_2$，则

$$T_{S_1 \times S_2}^0(x) = T_{S_1}^0(x_1) \times T_{S_2}^0(x_2),$$

$$N_{S_1 \times S_2}^0(x) = N_{S_1}^0(x_1) \times N_{S_2}^0(x_2).$$

证明 由定理 4.4.7 容易推得. □

定义 4.4.2 设集合 $S \subset \mathscr{B}$ 非空，点 $x^0 \in S$，$d \in \mathscr{B}$. 若对任给的 $\epsilon > 0$，存在 $t \in (0, \epsilon)$ 和 $p \in d + \epsilon B$ 使得 $x^0 + tp \in S$，则由所有这样的 d 组成的集合称为 S 在点 x^0 处的相依锥，记作

$Q_S(\boldsymbol{x}^0)$. 若有 $T_S^0(\boldsymbol{x}^0)=Q_S(\boldsymbol{x}^0)$，则称集合 S 在点 $\boldsymbol{x}^0$ 处是正则的.

注 4.4.3 设 $\boldsymbol{x}\in S$，据定理 4.4.7 易知有 $T_S^0(\boldsymbol{x})\subset Q_S(\boldsymbol{x})$，但反之不成立.

定理 4.4.9 设函数 $f:\mathscr{B}\to R$ 在 $\boldsymbol{x}\in\mathscr{B}$ 处是 Lipschitz 的，$\boldsymbol{0}\notin\partial^\circ f(\boldsymbol{x})$. 若 $S=\{\boldsymbol{y}\in\mathscr{B}\mid f(\boldsymbol{y})\leqslant f(\boldsymbol{x})\}$，则

$$\{\boldsymbol{d}\in\mathscr{B}\mid f^0(\boldsymbol{x};\boldsymbol{d})\leqslant 0\}\subset T_S^0(\boldsymbol{x}). \qquad (4.4.10)$$

再若 f 在 $\boldsymbol{x}$ 处是正则的，则集合 S 在 $\boldsymbol{x}$ 处是正则的，并且 (4.4.10)成为等式.

证明 设 $\boldsymbol{d}\in\mathscr{B}$ 使 $f^0(\boldsymbol{x};\boldsymbol{d})<0$，则存在 $\bar{\varepsilon},\bar{\delta}>0$，使得对任意的 $\boldsymbol{x}'\in\boldsymbol{x}+\bar{\varepsilon}B$ 和 $t\in(0,+\infty)$，有

$$f(\boldsymbol{x}'+t\boldsymbol{d})-f(\boldsymbol{x}')\leqslant-\bar{\delta}t.$$

设 $\{\boldsymbol{x}^k\}\subset S$，$\boldsymbol{x}^k\to\boldsymbol{x}$，$\{t_k\}\subset(0,+\infty)$，$t_k\to 0\ (k\to\infty)$，由 $f(\boldsymbol{x}^k)\leqslant f(\boldsymbol{x})$，故对任何 $k(=1,2,\cdots)$，有

$$f(\boldsymbol{x}^k+t_k\boldsymbol{d})\leqslant f(\boldsymbol{x}^k)-\bar{\delta}t_k\leqslant f(\boldsymbol{x})-\bar{\delta}t_k\leqslant f(\boldsymbol{x}),$$

从而 $\boldsymbol{x}^k+t_k\boldsymbol{d}\in S$. 由此，根据定理 4.4.7 得 $\boldsymbol{d}\in T_S^0(\boldsymbol{x})$. 因为 $\boldsymbol{0}\notin\partial^\circ f(\boldsymbol{x})$，则存在 $\boldsymbol{d}'\in S$，使得 $f^0(\boldsymbol{x};\boldsymbol{d}')<0$，即有 $\boldsymbol{d}'\in T_S^0(\boldsymbol{x})$. 设 $\boldsymbol{d}\in\{\boldsymbol{d}\in\mathscr{B}\mid f^0(\boldsymbol{x};\boldsymbol{d})\leqslant 0\}$，由定理 4.1.1 的(1)可知，对任意的 $\varepsilon>0$ 有

$$f^0(\boldsymbol{x};\boldsymbol{d}+\varepsilon\boldsymbol{d}')\leqslant f^0(\boldsymbol{x};\boldsymbol{d})+\varepsilon f^0(\boldsymbol{x};\boldsymbol{d}')<0,$$

从而

$$\boldsymbol{d}+\varepsilon\boldsymbol{d}'\in T_S^0(\boldsymbol{x}).$$

由 $T_S^0(\boldsymbol{x})$是闭的和 ϵ 的任意性，于是得 $\boldsymbol{d}\in T_S^0(\boldsymbol{x})$.

现设 $\boldsymbol{d}\in Q_S(\boldsymbol{x})$，由定义 4.4.2 有

$$\liminf_{t\to 0^+}\frac{d_S(\boldsymbol{x}+t\boldsymbol{d})}{t}=0,$$

即对任给的 $\epsilon>0$，存在收敛于 0 的正数列$\{t_k\}$，使得对充分大的 k

有 $d_S(x+t_kd)\leqslant \varepsilon t_k$，并存在点列 $\{x^k\}\subset S$ 使得

$$\|x+t_kd-x^k\|\leqslant 2\varepsilon t_k.$$

根据 f 在 x 处是 γ 秩 Lipschitz 的和 $f(x^k)\leqslant f(x)$，我们有

$$\frac{f(x+t_kd)-f(x)}{t_k}\leqslant\frac{f(x^k)+2\varepsilon\gamma t_k-f(x)}{t_k}\leqslant 2\varepsilon\gamma.$$

由于 f 在 x 处是正则的，在上式中令 $t_k\to 0$，$\varepsilon\to 0$ 可知 $f'(x;d)=f^\circ(x;d)\leqslant 0$，因此有 $\{d\in\mathscr{B}\,|\,f^\circ(x;d)\leqslant 0\}=T_S^\circ(x)$. ▯

推论 4.4.10 设条件与定理 4.4.9 相同，则

$$N_S^\circ(x)\subset\bigcup_{\lambda\geqslant 0}\lambda\partial^\circ f(x).$$

再若 f 在 x 处是正则的，则上式成为等式.

证明 由定理 4.4.2 和定理 4.4.9 可直接推得. ▯

推论 4.4.11 设 f_i: $\mathscr{B}\to R(i=1, 2, \cdots, m)$是实值函数，集合 $S=\{y\in\mathscr{B}\,|\,f_1(y)\leqslant 0, \cdots, f_m(y)\leqslant 0\}$，$x\in\mathscr{B}$ 使 $f_i(x)=0$ $(k=1, \cdots, m)$. 若 $f_i(i=1, \cdots, m)$ 在 x 处均是严格可微的，$D_Sf_1(x), \cdots, D_Sf_m(x)$是正线性独立的，并且 S 在 x 处是正则的，则

$$N_S^\circ(x)=\left\{\sum_{i=1}^m\lambda_iD_Sf_i(x)\,|\,\lambda_i\geqslant 0, i=1, \cdots, m\right\}.$$

证明 作函数 $f(y)=\max\{f_i(y)\,|\,i=1, \cdots, m\}$. 根据定理 4.2.12 可知，$f$ 在 x 处是 Lipschitz 的和正则的. 由 $S=\{y\in\mathscr{B}\,|\,f(y)\leqslant 0\}$ 和 $f(x)=0$，根据推论 4.4.10 即可推得结论. ▯

下面介绍另一个切锥的概念，并考虑它与 G-切锥之间的关系.

定义 4.4.3 设集合 $S\subset\mathscr{B}$ 非空，点 $x^0\in S$，$d\in\mathscr{B}$. 若对某个 $\varepsilon>0$，有 $p\in d+\varepsilon B$，$t\in(0, \varepsilon)$，使得

$$x'+tp\in S\quad\forall x'\in(x^0+\varepsilon B)\cap S,$$

则称 $\boldsymbol{d}$ 是集合 S 在 $\boldsymbol{x}^0$ 处的**超切向量**. S 在点 $\boldsymbol{x}^0$ 处的所有超切向量组成的集合称为 S 在点 $\boldsymbol{x}^0$ 处的**超切锥**,或 **H-切锥**,记作$T_S^H(\boldsymbol{x}^0)$.

注 4.4.4　设 $x \in \mathscr{B}$, 由定理 4.4.7 容易推得 $T_S^H(\boldsymbol{x}) \subset T_S^0(\boldsymbol{x})$.

定理 4.4.12　设集合 $S \subset \mathscr{B}$ 非空, $\boldsymbol{x} \in \mathscr{B}$. 若 $T_S^H(\boldsymbol{x}) \neq \varnothing$, 则

$$T_S^H(\boldsymbol{x}) = \mathrm{int} T_S^0(\boldsymbol{x}).$$

证明　首先证明下式成立:

$$T_S^H(\boldsymbol{x}) + T_S^0(\boldsymbol{x}) \subset T_S^H(\boldsymbol{x}). \tag{4.4.11}$$

事实上,设 $\boldsymbol{d}^1 \in T_S^H(\boldsymbol{x})$, $\boldsymbol{d}^2 \in T_S^0(\boldsymbol{x})$. 根据定义 4.4.3,存在 $\varepsilon_1 > 0$ 使得

$$S \cap (\boldsymbol{x} + \varepsilon_1 B) + t(\boldsymbol{d}^1 + \varepsilon_1 B) \subset S \quad \forall t \in (0, \varepsilon_1). \tag{4.4.12}$$

再由定理 4.4.7,存在 $\varepsilon_2 \leqslant 0$ 使得

$$S \cap (\boldsymbol{x} + \varepsilon_2 B) + t\boldsymbol{d}^2 \subset S + t(\varepsilon_1/2)B, \ t \in (0, \varepsilon_2). \tag{4.4.13}$$

取

$$\varepsilon < \min(\varepsilon_1/2, \ \varepsilon_1/(1 + \varepsilon_1 + \|\boldsymbol{d}^2\|), \ \varepsilon_2),$$

$$\boldsymbol{d} = \boldsymbol{y} + t(\boldsymbol{d}^1 + \boldsymbol{d}^2 + \varepsilon u),$$

其中 $\boldsymbol{y} \in S \cap (\boldsymbol{x} + \varepsilon B)$, $\boldsymbol{u} \in B$, $t \in (0, \varepsilon)$. 因为 $\varepsilon \leqslant \varepsilon_2$, 由 (4.4.13)知,对某个 $\boldsymbol{p} \in B$ 有

$$\boldsymbol{y} + t\boldsymbol{d}^2 - (1/2)t\varepsilon_1\boldsymbol{p} \in S.$$

另一方面,有

$$\|\boldsymbol{x} - \boldsymbol{y} - t\boldsymbol{d}^2 + (1/2)t\varepsilon_1\boldsymbol{p}\| \leqslant \varepsilon + \varepsilon(\|\boldsymbol{d}^2\| + \varepsilon_1) \leqslant \varepsilon_1.$$

从上面两式可得 $\boldsymbol{y} + t(\boldsymbol{d}^2 - (1/2)\varepsilon_1\boldsymbol{p}) \in S \cap (\boldsymbol{x} + \varepsilon_1 B)$. 由

(4.4.12)有

$$y + t\left(d^2 - (1/2)\epsilon_1 p\right) + t(d^1 + \epsilon_1 B) \subset S,$$

从而

$$d = y + t(d^2 - (1/2)\epsilon_1 p) + td^1$$

$$+ t(\epsilon u + (1/2)\epsilon_1 p) \subset S.$$

注意 $\|\epsilon u + (1/2)\epsilon_1 p\| \leqslant \epsilon + (1/2)\epsilon_1 < \epsilon_1$，根据定义 4.4.3 得 $d^1 + d^2 \in T_S^H(x)$，因此(4.4.11)成立.

显然，$T_S^H(x)$是开集，并且 $T_S^H(x) \subset T_S^0(x)$，故只要证明 $\mathrm{int}T_S^0(x) \subset T_S^H(x)$. 设 $d \in \mathrm{int}T_S^0(x)$，则对任意的 $p \in T_S^H(x)$，存在 $\bar{\lambda} > 0$ 使得 $d - \bar{\lambda}p \in \mathrm{int}T_S^0(x)$. 因为 $\bar{\lambda}p \in T_S^H(x)$，利用(4.4.11)，便得 $d = \bar{\lambda}p + (d - \bar{\lambda}p) \in T_S^H(x)$. ▯

推论 4.4.13 设集合 $S \subset \mathscr{B}$ 非空，$x \in S$. 若 $T_S^H(x) \neq \varnothing$，则对任何的 $\{x_k\} \subset \mathscr{B}$、$\{x_k^*\} \subset \mathscr{B}^*$，$x_k^* \in N_S^0(x_k)$，$x_k^* \to x^*\ (k \to \infty)$，有 $x^* \in N_S^0(x)$.

证明 设 $d \in T_S^H(x)$，则存在 $\bar{\epsilon} > 0$ 使得

$$S \cap (x + \bar{\epsilon}B) + t(d + \bar{\epsilon}B) \subset S \quad \forall t \in (0, \epsilon).$$

由此，对任意的 $y \in x + (\bar{\epsilon}/2)B$，有

$$S \cap \left(y + \frac{\bar{\epsilon}}{2}B\right) + t\left(d + \frac{\bar{\epsilon}}{2}B\right)$$

$$\subset S \cap (x + \bar{\epsilon}B) + t(d + \bar{\epsilon}B)$$

$$\subset S \quad \forall t \in \left(0, \frac{\bar{\epsilon}}{2}\right).$$

根据定义 4.4.3，对任意的 $y \in x + \frac{\bar{\epsilon}}{2}B$ 得 $d \in T_S^H(y)$，因此对任意的 $y \in x + \frac{\bar{\epsilon}}{2}B$ 有 $d \in T_S^0(y)$. 由 $x_k \to x$，$x_k^* \to x^*\ (k \to \infty)$，以及 $x_k^* \in N_S^0(x_k)$，有

$$\langle x^*, d\rangle = \lim_{k\to\infty}\langle x_k^*, d\rangle \leqslant 0 \quad \forall d \in T_S^H(x).$$

于是，由于定理 4.4.12 和 $T_S^0(x)$是闭的，得到对任意的 $d \in T_S^0(x)$ 有 $\langle x^*, d\rangle \leqslant 0$，从而 $x^* \in N_S^0(x)$. □

从以下的讨论可知，函数的上图象的 G-切锥和 G-法锥，分别与它的 G-方向导数和 G-次梯度有着等价关系. 因此，对于一般函数，我们可以用 G-切锥和 G-法锥来引进导数和次梯度. 由定义 2.1.6的(1)，函数 $f: \mathscr{B}\to R$ 的上图象为

$$\mathrm{epi} f = \{(x, \eta) \in \mathscr{B}\times R \mid \eta \geqslant f(x)\}.$$

定理 4.4.14 设函数 $f: \mathscr{B}\to R$ 在 $x \in \mathscr{B}$ 处是 Lipschitz 的.

(1) $(d, \eta) \in T_{\mathrm{epi}f}^0(x, f(x))$ 当且仅当 $f^0(x; d) \leqslant \eta$.

(2) f 在 x 处是正则的当且仅当集合 $\mathrm{epi}f$ 在$(x, f(x))$处是正则的.

证明 (1) 设 $(d, \eta) \in T_{\mathrm{epi}f}^0(x, f(x))$. 选取点列 $\{x^k\} \subset \mathscr{B}$，$x_k \to x$，$\{t_k\} \subset (0, +\infty)$，$t_k \to 0^+ (k\to\infty)$，使得

$$\lim_{k\to\infty}\frac{f(x^k + t_k d) - f(x^k)}{t_k} = f^0(x; d).$$

因为$(x^k, f(x^k)) \to (x, f(x))$，由定理 4.4.7 知存在序列 $\{d^k, \eta_k\} \subset \mathscr{B}\times R$，$(d^k, \eta_k) \to (d, \eta)$，使得 $\big(x^k, f(x^k)\big) + t_k(d^k, \eta_k) \in \mathrm{epi}f$. 因此，我们有

$$f(x^k) + t_k\eta_k \geqslant f(x^k + t_k d^k),\ k = 1, 2, \cdots,$$

即

$$\frac{f(x^k + t_k d) - f(x^k)}{t_k} \leqslant \eta_k,\ k = 1, 2, \cdots.$$

在上式中令 $k\to\infty$，得 $f^0(x; d) \leqslant \eta$.

设序列 $\{(x^k, \eta_k)\} \subset \mathrm{epi}f$，$x^k \to x$，$\eta_k \to f(x)$. 对任给的 $\delta > 0$，取

$$s_k = \max\left\{f^0(x; d) + \delta, \frac{f(x^k + t_k d) - f(x^k)}{t_k}\right\},$$

$$k = 1, 2, \cdots,$$

我们有

$$\eta_k + t_k s_k \geqslant \eta_k + f(\boldsymbol{x}^k + t_k \boldsymbol{d}) - f(\boldsymbol{x}^k),\ k = 1, 2, \cdots.$$

由于 $\eta_k \geqslant f(\boldsymbol{x}^k)$，故得

$$\eta_k + t_k s_k \geqslant f(\boldsymbol{x}^k + t_k \boldsymbol{d}),\ k = 1, 2, \cdots.$$

因为

$$\limsup_{k \to \infty} \frac{f(\boldsymbol{x}^k + t_k \boldsymbol{d}) - f(\boldsymbol{x}^k)}{t_k} \leqslant f^0(\boldsymbol{x}; \boldsymbol{d}),$$

所以当 $k \to \infty$ 时有 $s_k \to f^0(\boldsymbol{x}; \boldsymbol{d}) + \delta$. 由此，根据定理 4.4.7 得 $(\boldsymbol{d}, f^0(\boldsymbol{x}; \boldsymbol{d}) + \delta) \in T^0_{\mathrm{epi}f}(\boldsymbol{x}, f(\boldsymbol{x}))$，也就是当 $f^0(\boldsymbol{x}; \boldsymbol{d}) \leqslant \eta$ 时有 $(\boldsymbol{d}, \eta) \in T^0_{\mathrm{epi}f}(\boldsymbol{x}, f(\boldsymbol{x}))$.

(2) 若 f 在 $\boldsymbol{x}$ 处是正则的，作函数

$$g: \mathscr{B} \times R \to R,\ g(\boldsymbol{x}', \eta) = f(\boldsymbol{x}') - \eta.$$

显然，g 在 $(\boldsymbol{x}, f(\boldsymbol{x}))$ 处是正则的，并且 $\mathrm{epi}f = \{(\boldsymbol{x}, \eta) \mid g(\boldsymbol{x}, \eta) \leqslant 0\}$. 由于 $\boldsymbol{0} \notin \partial^0 g(\boldsymbol{x}, f(\boldsymbol{x})) = \partial^0 f(\boldsymbol{x}) \times \{-1\}$，由定理 4.4.9 可知集合 $\mathrm{epi}f$ 在 $(\boldsymbol{x}, f(\boldsymbol{x}))$ 处也是正则的.

反之，设

$$f'_+(\boldsymbol{x}; \boldsymbol{d}) = \liminf_{t \to 0^+} \frac{f(\boldsymbol{x} + t\boldsymbol{d}) - f(\boldsymbol{x})}{t},$$

我们证明 $Q_{\mathrm{epi}f}(\boldsymbol{x}, f(\boldsymbol{x})) = \mathrm{epi}f'_+(\boldsymbol{x}; \cdot)$ 成立. 事实上，设 $(\boldsymbol{d}, \eta) \in Q_{\mathrm{epi}f}(\boldsymbol{x}, f(\boldsymbol{x}))$，则对任给的 $\varepsilon > 0$，存在 $t \in (0, \varepsilon)$，$\boldsymbol{p} \in \boldsymbol{d} + \varepsilon B$ 和 $\eta' \in \eta + (-\varepsilon, \varepsilon)$ 使得 $(\boldsymbol{x}, f(\boldsymbol{x})) + t(\boldsymbol{p}, \eta') \in \mathrm{epi}f$，即 $f(\boldsymbol{x}) + t\eta' \geqslant f(\boldsymbol{x} + t\boldsymbol{p})$，从而有

$$\frac{f(\boldsymbol{x} + t\boldsymbol{p}) - f(\boldsymbol{x})}{t} \leqslant \eta' \leqslant \eta + \varepsilon.$$

令 $\varepsilon \to 0$，$t \to 0^+$，$\boldsymbol{p} \to \boldsymbol{d}$，对上式取极限，则得 $f'_+(\boldsymbol{x}; \boldsymbol{d}) \leqslant \eta$. 又对任

给的 $\delta > 0$，由 $f'_+(\boldsymbol{x};\boldsymbol{d})$的定义，则必存在 $t_k \to 0^+$，使得

$$\frac{f(\boldsymbol{x}+t_k\boldsymbol{d})-f(\boldsymbol{x})}{t_k} \leqslant f'_+(\boldsymbol{x};\boldsymbol{d})+\delta,\ k=1,2,\cdots.$$

取 $s = f'_+(\boldsymbol{x};\boldsymbol{d})+\delta$，有

$$f(\boldsymbol{x})+t_k s \geqslant f(\boldsymbol{x}+t_k\boldsymbol{d}),\ k=1,2,\cdots.$$

由此，据定义 4.4.2 知 $(\boldsymbol{d}, f'_+(\boldsymbol{x};\boldsymbol{d})+\delta) \in Q_{\mathrm{epi}f}(\boldsymbol{x}, f(\boldsymbol{x}))$. 由于 $\mathrm{epi}f$ 在$(\boldsymbol{x}, f(\boldsymbol{x}))$处是正则的，由(1)的证明我们有

$$\mathrm{epi}f'_+(\boldsymbol{x};\ \cdot\ ) = \mathrm{epi}f^0(\boldsymbol{x};\ \cdot\ ).$$

于是，对任意的 $\boldsymbol{d} \in \mathscr{B}$ 有 $f'_+(\boldsymbol{x};\boldsymbol{d}) = f^0(\boldsymbol{x};\boldsymbol{d})$，从而 $f'(\boldsymbol{x};\boldsymbol{d}) = f^0(\boldsymbol{x};\boldsymbol{d})$. 据此，由定义 4.2.4 得 f 在 $\boldsymbol{x}$ 处是正则的. □

推论 4.4.15　设 $f:\mathscr{B}\to R$ 是实值函数，$\boldsymbol{x}\in\mathscr{B}$，则 $\boldsymbol{x}^* \in \partial^0 f(\boldsymbol{x})$ 当且仅当 $(\boldsymbol{x}^*, -1) \in N^0_{\mathrm{epi}f}(\boldsymbol{x}, f(\boldsymbol{x}))$.

证明　因为 $\boldsymbol{x}^* \in \partial^0 f(\boldsymbol{x})$ 当且仅当对任意的 $\boldsymbol{d}\in\mathscr{B}$ 有 $f^0(\boldsymbol{x};\boldsymbol{d}) \geqslant \langle \boldsymbol{x}^*, \boldsymbol{d}\rangle$. 这意味着对任意的 $\boldsymbol{d}\in\mathscr{B}$ 和任何 $\eta \geqslant f^0(\boldsymbol{x};\boldsymbol{d})$，有

$$\langle(\boldsymbol{x}^*, -1), (\boldsymbol{d}, \eta)\rangle \leqslant 0.$$

根据定理 4.4.5，它等价于 $(\boldsymbol{d}, \eta) \in T^0_{\mathrm{epi}f}(\boldsymbol{x}, f(\boldsymbol{x}))$，由定义 4.4.1 即得 $(\boldsymbol{x}^*, -1) \in N^0_{\mathrm{epi}f}(\boldsymbol{x}, f(\boldsymbol{x}))$. □

推论 4.4.15 告诉我们，可以利用函数上图象的 G-法锥来定义它的 G-次梯度，并且该函数可以不是 Lipschitz 的. 我们将在第 5 章进一步研究由切锥引进次梯度的问题.

利用 G-法锥，可以给出一般函数的 G-次梯度的定义如下.

定义 4.4.4　设函数 $f:\mathscr{B}\to R\cup\{\pm\infty\}$ 在点 $\boldsymbol{x}^0\in\mathscr{B}$ 处是有限的. 若存在 $\boldsymbol{x}^* \in \mathscr{B}^*$ 使得 $(\boldsymbol{x}^*, -1) \in N^0_{\mathrm{epi}f}(\boldsymbol{x}^0, f(\boldsymbol{x}^0))$，则称 $\boldsymbol{x}^*$ 是 f 在点 $\boldsymbol{x}^0$ 处的 *G-次梯度*. f 在点 $\boldsymbol{x}^0$ 处的所有 G-次梯度的集合称为 f 在点 $\boldsymbol{x}^0$ 处的 *G-次微分*，记作 $\partial^0 f(\boldsymbol{x}^0)$. 若集合 $\mathrm{epi}f$ 在点$(\boldsymbol{x}^0, f(\boldsymbol{x}^0))$处是正则的，则称 f 在点 $\boldsymbol{x}^0$ 处是*正则的*.

定理 4.4.16　设函数 $f:\mathscr{B}\to R\cup\{\pm\infty\}$ 在 $\bar{\boldsymbol{x}}\in\mathscr{B}$ 处有限.

若对 $\tilde{x}$ 的某一邻域内任何一点 x 有 $f(\tilde{x}) \leqslant f(x)$，则 $\boldsymbol{0} \in \partial^0 f(\tilde{x})$.

证明 对任意的 $(\boldsymbol{d}, \eta) \in T^0_{\mathrm{epi}f}(\tilde{x}, f(\tilde{x}))$，存在序列 $\{(\boldsymbol{d}^k, \eta_k)\}$，$(\boldsymbol{d}^k, \eta_k) \to (\boldsymbol{d}, \eta)$ 和 $\{t_k\}$，$t_k \to 0^+ (k \to \infty)$，使得 $(\tilde{x}, f(\tilde{x})) + t_k(\boldsymbol{d}^k, \eta_k) \in \mathrm{epi}f$，即有

$$f(\tilde{x}) + t_k\eta_k \geqslant f(\tilde{x}, t_k\boldsymbol{d}^k).$$

由此，对充分大的 k，我们有 $f(\tilde{x}) + t_k\eta_k \geqslant f(\tilde{x}, t_k\boldsymbol{d}^k) \geqslant f(\tilde{x})$，从而 $t_k\eta_k \geqslant 0$. 由 $t_k > 0$，则有 $\lim\limits_{k\to\infty}\eta_k = \eta \geqslant 0$，于是对任意的 $(\boldsymbol{d}, \eta) \in T^0_{\mathrm{epi}f}(\tilde{x}, f(\tilde{x}))$，有 $\langle(\boldsymbol{0}, -1), (\boldsymbol{d}, \eta)\rangle \leqslant 0$，故 $\boldsymbol{0} \in \partial^0 f(\tilde{x})$. ▯

定理 4.4.17 设集合 $S \subset \mathscr{B}$ 非空. 对于 S 上的指示函数

$$\delta_S(x) = \begin{cases} 0, & x \in S; \\ +\infty, & x \in \mathscr{B} \backslash S, \end{cases}$$

若 $x \in S$，则 $\partial^0 \delta_S(x) = N^0_S(x)$. 并且，函数 δ_S 在 x 处是正则的当且仅当 S 在 x 处是正则的.

证明 $x^* \in \partial^0\delta_S(x)$ 等价于 $(x^*, -1) \in N^0_{\mathrm{epi}g}(x, \boldsymbol{0})$. 显然有 $\mathrm{epi}\delta_S = S \times [0, +\infty)$. 根据推论 4.4.8，它等价于 $x^* \in N^0_S(x)$，$-1 \in N^0_{R_+}(\boldsymbol{0})$. ▯

§4.5 有限维情况

本节讨论当 $\mathscr{B}$ 为 Euclid 空间 R^n 时，G-次梯度和 G-次微分所特有的性质. 这时有 $\mathscr{B}^* = R^n$，以及 $\partial^0 f(x) \subset R^n$.

一个重要的结果是，Lipschitz 函数在 R^n 的开子集上是几乎处处可微的. 由此，函数在某点处的 G-次梯度可以由其附近的可微点逼近. 又几乎处处可微的点在 G-次可微点集上是稠密的. 以下记 Ω_f 表示 f 的不是 G-次可微的点组成的集合.

定理 4.5.1 设函数 $f: R^n \to R$ 在 $x \in R^n$ 处是 Lipschitz 的，$S \subset R^n$ 是 Lebesgue 测度为 0 的集合，则有

$$\partial^0 f(\boldsymbol{x}) = \operatorname{co}\{\lim_{k\to\infty} \nabla f(\boldsymbol{x}^k) \mid \boldsymbol{x}^k \to \boldsymbol{x},\ \boldsymbol{x}^k \notin S,\ \boldsymbol{x}^k \notin \Omega_f\}.$$

证明　设点列 $\{\boldsymbol{x}^k\} \subset R^n$, $\boldsymbol{x}^k \to \boldsymbol{x}\ (k \to \infty)$, $\boldsymbol{x}^k \notin S \cap \Omega_f$, $S \cup \Omega_f$ 的测度为 0, f 在 $\boldsymbol{x}^k$ 处 G-次可微. 由定理 4.1.1 可知 $\partial^0 f$ 在 x附近是局部有界的,又由定理 4.2.5,则有 $\nabla f(\boldsymbol{x}^k) \in \partial^0 f(\boldsymbol{x}^k)$ $(k = 1, 2, \cdots)$. 根据 Bolzano-Weierstrass 定理,序列 $\{\nabla f(\boldsymbol{x}^k)\}$ 存在收敛的子列,因为 $\partial^0 f(\boldsymbol{x})$是闭的,所以$\{\nabla f(\boldsymbol{x}^k)\}$的聚点均属于 $\partial^0 f(\boldsymbol{x})$,即

$$\{\lim_{k\to\infty} \nabla f(\boldsymbol{x}^k) \mid \boldsymbol{x}^k \to \boldsymbol{x},\ \boldsymbol{x}^k \notin S \cup \Omega_f\} \subset \partial^0 f(\boldsymbol{x}).$$

由于 $\partial^0 f(\boldsymbol{x})$是凸集,因此上式左边的凸包也包含于 $\partial^0 f(\boldsymbol{x})$. 下面证明对任意的 $\varepsilon > 0$ 和任意的 $\boldsymbol{d}(\in R^n) \neq \boldsymbol{0}$, 有

$$f^0(\boldsymbol{x};\ \boldsymbol{d}) - \varepsilon \leqslant \limsup_{y\to x} \{\nabla f(\boldsymbol{y})^T \boldsymbol{d} \mid \boldsymbol{y} \to \boldsymbol{x},\ \boldsymbol{y} \notin S \cup \Omega_f\}.$$

为此,设 $a = \limsup_{k\to\infty} \{\nabla f(\boldsymbol{y})^T \boldsymbol{d} \mid \boldsymbol{y} \to \boldsymbol{x},\ \boldsymbol{y} \notin S \cup \Omega_f\}$. 由 a 的定义则存在 $\delta > 0$, 使得

$$\boldsymbol{y} \in \boldsymbol{x} + \delta B,\ \boldsymbol{y} \notin S \cup \Omega_f,\ \nabla f(\boldsymbol{y})^T \boldsymbol{d} \leqslant a + \varepsilon. \tag{4.5.1}$$

可选 δ 充分小使得 $\mu((S \cup \Omega_f) \cap (\boldsymbol{x} + \delta B)) = 0$($\mu(\cdot)$ 是 R^n 中的测度函数),并考虑线段

$$L_y = \left\{\boldsymbol{y} + t\boldsymbol{d} \,\middle|\, 0 < t < \frac{\delta}{2\|\boldsymbol{d}\|}\right\}.$$

根据 Fubini 定理[32],对于任意的 $\boldsymbol{y} \in \boldsymbol{x} + (\delta/2)B$, $\boldsymbol{y} \notin S \cup \Omega_f$ 和 $L_y \subset \boldsymbol{x} + (\delta/2)B + t\boldsymbol{d} \subset \boldsymbol{x} + \delta B$, 有 $\mu((S \cup \Omega_f) \cap L_y) = 0$. 设 $t \in \left(0, \frac{\delta}{2\|\boldsymbol{d}\|}\right)$, 则有

$$f(\boldsymbol{y} + t\boldsymbol{d}) - f(\boldsymbol{y}) = \int_0^t \nabla f(\boldsymbol{y} + s\boldsymbol{d})^T \boldsymbol{d}(\mathrm{d}s).$$

因为 $\|\boldsymbol{y} + s\boldsymbol{d} - \boldsymbol{x}\| \leqslant \delta$, $0 < s < t$, 由(4.5.1)和上式可得

$$f(y+td)-f(y)\leqslant t(a+\varepsilon)$$

$$\forall y\in x+(\delta/2)B,\ y\notin S\cup\Omega_f,$$

再由 f 是连续的，可知上式对任意的 $y\in x+(\delta/2)B$、$t\in\left(0,\frac{\delta}{2\|d\|}\right)$ 均成立，故有 $f^{\circ}(x;d)\leqslant a+\epsilon$. 由此，推得定理的结论成立. ▯

显然有以下推论.

推论 4.5.2 设函数 $f: R^n\to R$ 在 $x\in R^n$ 处是 Lipschitz 的，则

$$f^{\circ}(x;d)=\limsup_{y\to x}\{\nabla f(y)^T d\mid y\notin S\cup\Omega_f\}.$$

例 4.5.1 考虑函数 $f: R^2\to R$, $f(x,y)=\max\{\min\{x,-y\},y-x\}$. 设

$$S_1=\{(x,y)\mid y\leqslant 2x,\ y\leqslant -x\},$$

$$S_2=\{(x,y)\mid y\leqslant 2x,\ y\geqslant -x\}$$

和

$$S_3=\{(x,y)\mid y\geqslant 2x,\text{或 } y\geqslant x/2\},$$

则 $S_1\cup S_2\cup S_3=R^2$, $f(x,y)$等价为

$$f(x,y)=\begin{cases}x, & (x,y)\in S_1;\\ -y, & (x,y)\in S_2;\\ y-x, & (x,y)\notin S_3.\end{cases}$$

设 S 是 S_1、S_2 和 S_3 的边界构成的集合，S 的测度为 0. 对于 $(x,y)\notin S$ 的点，f 是可微的，并且$\nabla f(x,y)$是$(1,0)$或$(0,-1)$或$(-1,1)$，故 $\partial^{\circ}f(0,0)$是这三个点的凸包(图 4.5.1). 由 $f(x,0)=\max\{0,-x\}$，则 $\partial_1^{\circ}f(0,0)=[-1,0]$，又 $f(0,y)=\max\{0,y\}$，故 $\partial_2^{\circ}f(0,0)=[0,1]$. 于是有

$$\partial^{\circ}f(0,0)\neq\partial_1^{\circ}f(0,0)\times\partial_2^{\circ}f(0,0).$$

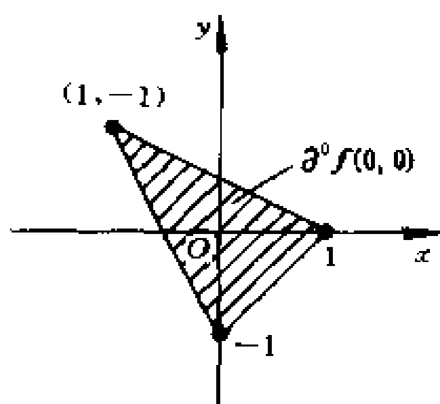


图　4.5.1

定理 4.5.3　设 $\boldsymbol{a} \in R^n$, $\boldsymbol{b} \in R^m$, 函数 $f: R^n \times R^m \to R$ 在点 $\boldsymbol{x} = (\boldsymbol{a}, \boldsymbol{b})$ 的凸邻域 $U \times V \subset R^n \times R^m$ 上是 Lipschitz 的，对于每个 $\boldsymbol{a}' \in U$，函数 $f(\boldsymbol{a}', \cdot)$ 在 V 上是凸的. 若 $(\boldsymbol{z}, \boldsymbol{w}) \in \partial^0 f(\boldsymbol{x})$，则 $\boldsymbol{w} \in \partial_2^0 f(\boldsymbol{a}, \boldsymbol{b})$.

证明　由定理 4.5.1，对于 $(\boldsymbol{z}, \boldsymbol{w}) \in \partial^0 f(\boldsymbol{x})$，存在 $\boldsymbol{x}^k = (\boldsymbol{a}^k, \boldsymbol{b}^k) \to \boldsymbol{x}$ $(k \to \infty)$, $\boldsymbol{x}^k \notin \Omega_f$，使得 $\nabla f(\boldsymbol{x}^k) \to (\boldsymbol{z}, \boldsymbol{w})$. 设 $\nabla f(\boldsymbol{x}^k) = (\boldsymbol{z}^k, \boldsymbol{w}^k)$，根据定理 4.2.6 和定理 4.2.11 知 $\boldsymbol{w}^k \in \partial_2^0 f(\boldsymbol{a}^k, \boldsymbol{b}^k)$. 因为 $f(\boldsymbol{a}^k, \cdot)$ 关于 $\boldsymbol{b}^k$ 是凸函数，故对任意的 $\boldsymbol{d} \in R^m$ 和充分大的 k，有

$$f(\boldsymbol{a}^k, \boldsymbol{b}^k + \boldsymbol{d}) - f(\boldsymbol{a}^k, \boldsymbol{b}^k) \geqslant \langle \boldsymbol{d}, \boldsymbol{w}^k \rangle.$$

对上式令 $k \to \infty$，我们有

$$f(\boldsymbol{a}, \boldsymbol{b} + \boldsymbol{d}) - f(\boldsymbol{a}, \boldsymbol{b}) \geqslant \langle \boldsymbol{d}, \boldsymbol{w} \rangle.$$

由此，即得 $\boldsymbol{w} \in \partial_2^0 f(\boldsymbol{a}, \boldsymbol{b})$. □

考虑 R^n 中集合 S 的距离函数：

$$d_S(\boldsymbol{x}) = \inf_{y \in S} \| \boldsymbol{x} - \boldsymbol{y} \|, \quad \boldsymbol{x} \in R^n,$$

其中 $\| \boldsymbol{x} - \boldsymbol{y} \| = \left(\sum_{i=1}^n |x_i - y_i|^2 \right)^{\frac{1}{2}}$, $\boldsymbol{x} = (x_1, \cdots, n)^T$, $\boldsymbol{y} = (y_1, \cdots, y_n)^T$.

定理 4.5.4 设集合 $S \subset R^n$ 非空. 若 $\nabla d_S(\boldsymbol{x})$ 存在且不为 $\boldsymbol{0}$, 则 $\boldsymbol{x} \notin \mathrm{cl}S$, 并且存在唯一的点 $\boldsymbol{x}^0 \in \mathrm{cl}S$, 使得

$$\nabla d_S(\boldsymbol{x}) = \frac{\boldsymbol{x} - \boldsymbol{x}^0}{\|\boldsymbol{x} - \boldsymbol{x}^0\|}.$$

证明 反证:假设 $\boldsymbol{x} \in \mathrm{cl}S$, 则对任意的 $\boldsymbol{d} \in R^N$ 有

$$\nabla d_S(\boldsymbol{x})^T \boldsymbol{d} = \lim_{t \to 0^+} \frac{d_S(\boldsymbol{x} + t\boldsymbol{d}) - d_S(\boldsymbol{x})}{t}$$

$$= \lim_{t \to 0^-} \frac{d_S(\boldsymbol{x} + t\boldsymbol{d})}{t} \geqslant 0.$$

由此得 $\nabla d_S(\boldsymbol{x}) = \boldsymbol{0}$, 它与已知相矛盾. 从 $\boldsymbol{x} \notin \mathrm{cl}S$, 则存在一点 $\boldsymbol{x}^0 \in \mathrm{cl}S$ 使得 $d_S(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{x}^0\|$. 设 $t \in (0, 1)$, 则显然有

$$d_S\big(\boldsymbol{x} + t(\boldsymbol{x}^0 - \boldsymbol{x})\big) = (1 - t)\|\boldsymbol{x} - \boldsymbol{x}^0\|,$$

于是

$$\frac{d_S\big(\boldsymbol{x} + t(\boldsymbol{x}^0 - \boldsymbol{x})\big) - d_S(\boldsymbol{x})}{t} = -\|\boldsymbol{x} - \boldsymbol{x}^0\|.$$

令 $t \to 0^+$, 得到

$$d_S'(\boldsymbol{x}; \boldsymbol{x}^0 - \boldsymbol{x}) = \nabla d_S(\boldsymbol{x})^T(\boldsymbol{x}^0 - \boldsymbol{x}) = -\|\boldsymbol{x} - \boldsymbol{x}^0\|.$$

因为 d_S 是 1 秩 Lipschitz 的, 所以 $\|\nabla d_S(\boldsymbol{x})\| \leqslant 1$, 从而有 $\nabla d_S(\boldsymbol{x}) = \dfrac{\boldsymbol{x} - \boldsymbol{x}^0}{\|\boldsymbol{x} - \boldsymbol{x}^0\|}$. □

定义 4.5.1 设集合 $S \subset \mathscr{B}$ 非空, $\boldsymbol{d} \in \mathscr{B}$, $\boldsymbol{d} \neq \boldsymbol{0}$. 若存在唯一的点 $\boldsymbol{x}^0 \in \mathrm{cl}S$, 使得 $\boldsymbol{d} = \boldsymbol{x} - \boldsymbol{x}^0$, 其中 $\boldsymbol{x} \in \mathscr{B}$ 满足 $d_S(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{x}^0\|$, 则称 $\boldsymbol{d}$ 在点 $\boldsymbol{x}^0$ 处垂直于 S, 记作 $d \perp S$ (在 $\boldsymbol{x}^0$ 处).

定理 4.5.5 设集合 $S \subset R^n$ 非空, $\boldsymbol{d} \in R^n$, $\boldsymbol{d} \neq \boldsymbol{0}$. 若 $\boldsymbol{d}$ 在 $\boldsymbol{x} \in \mathrm{cl}S$ 处垂直于 S, 则对任意的 $\boldsymbol{y} \in \mathrm{cl}S$, 有

$$\langle \boldsymbol{d}, \boldsymbol{y} - \boldsymbol{x} \rangle \leqslant \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{x}\|^2.$$

证明　由定义 4.5.1，对任意的 $\boldsymbol{y} \in \mathrm{cl}S$，有 $\boldsymbol{d} = \boldsymbol{x}' - \boldsymbol{x}$ 和

$$\| \boldsymbol{x}' - \boldsymbol{y} \| \geqslant \| \boldsymbol{x}' - \boldsymbol{x} \| \quad \forall \boldsymbol{x}' \in S.$$

从上式有

$$\langle \boldsymbol{x}' - \boldsymbol{y}, \boldsymbol{x}' - \boldsymbol{y} \rangle \geqslant \langle \boldsymbol{x}' - \boldsymbol{x}, \boldsymbol{x}' - \boldsymbol{x} \rangle = \| \boldsymbol{d} \|^2.$$

据此，得

$$\begin{aligned} & \| \boldsymbol{d} \|^2 + 2\langle \boldsymbol{d}, \boldsymbol{x} - \boldsymbol{y} \rangle + \| \boldsymbol{x} - \boldsymbol{y} \|^2 \\ & = \langle \boldsymbol{d} + \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{d} + \boldsymbol{x} - \boldsymbol{y} \rangle \\ & = \langle \boldsymbol{x}' - \boldsymbol{y}, \boldsymbol{x}' - \boldsymbol{y} \rangle \geqslant \| \boldsymbol{d} \|^2, \end{aligned}$$

所以结论成立. ▯

定理 4.5.6　设集合 $S \subset R^n$ 非空，$\boldsymbol{x} \in \mathrm{cl}S$，则

$$\partial^\circ d_S(\boldsymbol{x}) = \mathrm{co}\left(\{\boldsymbol{0}\} \cup \left\{ \boldsymbol{d} = \lim_{k\to\infty} \frac{\boldsymbol{d}^k}{\| \boldsymbol{d}^k \|} \,\middle|\, \boldsymbol{d}^k \perp S (\text{在 } \boldsymbol{x}^k \text{ 处}),\ \boldsymbol{x}^k \to \boldsymbol{x},\ \boldsymbol{d}^k \to \boldsymbol{0} \right\} \right).$$

证明　设

$$A = \mathrm{co}\left(\{\boldsymbol{0}\} \cup \left\{ \boldsymbol{d} = \lim_{k\to\infty} \frac{\boldsymbol{d}^k}{\| \boldsymbol{d}^k \|} \,\middle|\, \boldsymbol{d}^k \perp S (\text{在 } \boldsymbol{x}^k \text{ 处}),\ \boldsymbol{x}^k \to \boldsymbol{x},\ \boldsymbol{d}^k \to \boldsymbol{0} \right\} \right),$$

由定理 4.5.4 知 $\nabla d_S(\boldsymbol{x})$ 或为 $\boldsymbol{0}$ 或为单位向量. 因此，根据定理 4.5.1 得知 $\partial^\circ d_S(\boldsymbol{x}) \subset A$.

显然，d_S 在 $\boldsymbol{x}$ 处达到极小值，故 $\boldsymbol{0} \in \partial^\circ d_S(\boldsymbol{x})$. 下设 $\boldsymbol{d}^k$ 在 $\boldsymbol{x}^k$ 处垂直于 S，即 $\boldsymbol{d}^k = \boldsymbol{y}^k - \boldsymbol{x}^k$，$d_S(\boldsymbol{y}^k) = \| \boldsymbol{y}^k - \boldsymbol{x}^k \|$，并且 $\boldsymbol{x}^k \to \boldsymbol{x}$，$\boldsymbol{d}^k \to \boldsymbol{0}$. 利用定理 4.3.5，我们有

$$d_S(\boldsymbol{y}^k) - d_S(\boldsymbol{x}^k) \in \langle \partial^\circ d_S(\boldsymbol{\xi}^k), \boldsymbol{y}^k - \boldsymbol{x}^k \rangle,$$

其中 $\xi^k \in (x^k, y^k)$. 因为 $d_S(x^k) = 0$, 上式可写成

$$1 \in \left\langle \partial^0 d_S(\xi^k), \frac{y^k - x^k}{\| y^k - x^k \|} \right\rangle.$$

又因为 $\| \partial^0 d_S(x) \| \leqslant 1$, 从而可得 $\frac{d^k}{\| d^k \|} \in \partial^0 d_S(\xi^k)$. 由 $x^k \to x$, $d^k \to 0$, 有 $\xi^k \to x$, 依据定理 4.2.1 的(2)得 $d \in \partial^0 d_S(x)$, 因此 $A \subset \partial^0 d_S(x)$. ▯

推论 4.5.7 设集合 $S \subset R^n$ 非空. 若 x 属于 $\mathrm{cl}S$ 的边界点,则 $\partial^0 d_S(x)$包含非零点.

证明 因 x 属于 $\mathrm{cl}S$ 的边界点,则对任意的 $\epsilon > 0$, 存在点 $y_\epsilon \in R^n$, 使得 $\| y_\epsilon - x \| \leqslant \epsilon$, 并且 $y_\epsilon \notin \mathrm{cl}S$. 由定理 4.5.4 知,存在垂直向量 d_ϵ 使得 $\nabla d_S(y_\epsilon) = \frac{d_\epsilon}{\| d_\epsilon \|}$. 令 $\epsilon \to 0^+$, 有 $y_\epsilon \to x$, 则 $\frac{d_\epsilon}{\| d_\epsilon \|} \to d \in \partial^0 d_S(x)$, $\| d \| = 1$. ▯

推论 4.5.8 设集合 $S \subset R^n$ 非空. 若 x 属于 $\mathrm{cl}S$ 的边界点,则 $N_S^0(x)$包含非零点.

证明 由定理 4.4.2 和推论 4.5.7 即得证. ▯

定理 4.5.9 设集合 $S \subset R^n$ 非空, $x \in \mathrm{cl}S$, 并且

$$A = \mathrm{co}\left\{ \{0\} \cup \left\{ d = \lim_{k \to \infty} \frac{d^k}{\| d^k \|} \,\middle|\, d^k \perp S(\text{在 } x^k \text{ 处}), x^k \to x,\ d^k \to 0 \right\} \right\}.$$

则 $N_S^0(x) = \bigcup_{\lambda \geqslant 0} \lambda A$.

证明 根据定理 4.4.2, $N_S^0(x)$是闭凸锥,再由定理 4.5.6 即可得证. ▯

定理 4.5.10 设 $S \subset R^n$ 是闭集, $x \in S$, 则 $d \in \mathrm{int} T_S^0(x)$ 当且仅当存在 $\varepsilon > 0$ 使得

$$d_S(y + tp) \leqslant d_S(y) \quad \forall y \in x + \varepsilon B,$$

$$p \in d + \varepsilon B,\ t \in [0, \varepsilon). \tag{4.5.2}$$

证明　若存在 d 满足(4.5.2),则对任意的 $p \in d + \varepsilon B$, $x \in S$, 可以找到点列 $\{x^k\} \subset (x + \varepsilon B) \cap S$, $x^k \to x$, $\{t_k\} \in (0, \varepsilon)$, $t_k \to 0^+ (k \to \infty)$, 对充分大的 k 有

$$d_S(x^k + t_k p) \leqslant d_S(x^k) = 0.$$

因此,对充分大的 k 我们有 $x^k + t_k p \in S$. 再由定理 4.4.7 知 $p \in T_S^0(x)$, 故得 $d \in \mathrm{int} T_S^0(x)$.

设 $d \in \mathrm{int} T_S^0(x)$, 则对任一非零的 $z \in N_S^0(x)$ 有 $d^T z < 0$(否则,若 $d^T z = 0$, 则在 d 的任一邻域中存在一点 $\bar{p}$ 使得 $\bar{p}^T z > 0$, 但另一方面,从 $p \in T_S^0(x)$ 有 $p^T z \leqslant 0$, 导致矛盾). 因此,存在 $\gamma > 0$ 使得

$$p^T z \leqslant -\gamma \| z \| \quad \forall z \in N_S^0(x).$$

利用定理 4.4.2,从上式可推知

$$p^T z \leqslant -\gamma \| z \| \quad \forall z \in \partial^0 d_S(x). \tag{4.5.3}$$

现在证明:存在某个 $\bar{w} > 0$, 使得对任意的 $p \in d + \bar{w} B$ 有

$$p^T \nabla d_S(y) \leqslant 0 \quad \forall y \in x + \bar{w} B. \tag{4.5.4}$$

反之,假若(4.5.4)不成立,则存在点列 $\{y^k\} \subset R^n$, $y^k \to x (k \to \infty)$, 使得

$$p^T \nabla d_S(y^k) > 0.$$

由定理 4.5.4, $\nabla d_S(y^k)$ 是非零的单位向量,当 $k \to \infty$ 时,它收敛到点 $z \in \partial^0 d_S(x)$. 对上式取极限得

$$p^T z \geqslant 0,$$

导致与(4.5.3)矛盾.

当 $y \in x + \varepsilon B$, $p \in d + \varepsilon B$, $t \in [0, \varepsilon)$ 时,选取 ε 使得 $y + tp \in x + wB$. 注意,函数 d_S 是几乎处处可微的,固定 p, 有

$$d_S(y+tp)-\delta_S(y)=\int_0^t \nabla d_S(y+sp)^T p ds.$$

最后，利用(4.5.4)，对几乎所有的 $y \in x+\epsilon B$，有 $d_S(y+tp) \leqslant d_S(y)$. 由 d_S 的连续性知(4.5.2)成立. ▯

推论 4.5.11 设 $S \subset R^n$ 是闭集，$x \in S$. 则 $T_S^H(x) = \mathrm{int} T_S^0(x)$.

证明 根据定理 4.5.10 知，由于 $d \in T_S^H(x)$，从(4.5.2)有 $d \in \mathrm{int} T_S^0(x)$. 反之，设 $d \in \mathrm{int} T_S^0(x)$，由定理 4.4.12 得 $d \in T_S^H(x)$. ▯

推论 4.5.12 设 $S \subset R^n$ 是闭集，$x \in S$. 若 $\mathrm{int} T_S^0(x) \neq \varnothing$，则对任意的 $x^k \to x$，$\zeta^k \in N_S(x^k)$，$\zeta^k \to \zeta\ (k \to \infty)$，有 $\zeta \in N_S^0(x)$.

证明 由推论 4.5.11 容易推得. ▯

§ 4.6 Lipschitz 向量函数的 G-次微分

最后，考虑满足 Lipschitz 条件的向量函数的 G-次微分问题.

定义 4.6.1 设集合 $S \subset R^n$ 非空，$f: X \to R^m$ 是向量函数 $(m \geqslant 1)$，$f(x) = (f_1(x), \cdots, f_m(x))^T$. 若每一个 $f_i(i=1,\cdots,m)$ 在点 $x^0 \in S$ 处是(γ 秩)Lipschitz 的，则称向量函数 f 在点 x^0 处是 Lipschitz 的. 若每一个 $f_i(i=1,\cdots,m)$ 是 S 上的(γ 秩)Lipschitz 函数，则称 f 是 S 上的(γ 秩)Lipschitz **向量函数**.

设 $f(x) = (f_1(x), \cdots, f_m(x))^T$ 在点 $x \in R^n$ 处是 Lipschitz 的，这时每一个 $f_i(i=1,\cdots,m)$ 是几乎处处可微的. 相应地，我们说 f 在点 x 的邻域内是几乎处处可微的. 用 Ω_f 表示 f 的不可微点组成的集合. 若各 $f_i(i=1,\cdots,m)$ 在 x 处的偏导数均存在，则用 $Jf(x)$表示由这些偏导数构成的 $m \times n$ 阶 Jacobi 矩阵.

定义 4.6.2 设 $x^0 \in R^n$，$x^k \in R^n(k=1,2,\cdots)$，$f: R^n \to R^m$ 是 Lipschitz 向量函数. 记

$$\partial^0 f(x^0) = \mathrm{co}\left\{\lim_{k\to\infty} Jf(x^k) \,|\, x^k \to x^0,\ x^k \notin \Omega_f\right\},$$

则称 $\boldsymbol{x}^* \in \partial^0 \boldsymbol{f}(\boldsymbol{x}^0)$ 是 $\boldsymbol{f}$ 在点 $\boldsymbol{x}^0$ 处的 *G-次梯度*，称 $\partial^0 \boldsymbol{f}(\boldsymbol{x}^0)$是 $\boldsymbol{f}$ 在点 $\boldsymbol{x}^0$ 处的 *G-次微分*. 若 $\partial^0 \boldsymbol{f}(\boldsymbol{x}^0) \neq \varnothing$，则称 $\boldsymbol{f}$ 在点 $\boldsymbol{x}^0$ 处是 *G-次可微的*.

对于 $m \times n$ 阶矩阵 $M \in R^{m\times n}$，定义它的范数：

$$\| M \|_{m\times n} = \left\{ \sum_{i=1}^{m} \| r_i \|^2 \,|\, r_i \text{ 是 } M \text{ 的第 } i \text{ 行向量} \right\}^{1/2},$$

并记 $B_{m\times n} \in R^{m\times n}$ 表示开单位球.

定理 4.6.1 设 $\boldsymbol{f}: R^n \to R^m$ 在 $\boldsymbol{x} \in R^n$ 处是 Lipschitz 的，$\boldsymbol{f}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), \cdots, f_m(\boldsymbol{x}))^T$.

(1) $\partial^0 \boldsymbol{f}(\boldsymbol{x}) \subset R^{m\times n}$ 是非空紧凸集.

(2) $\partial^0 \boldsymbol{f}(\boldsymbol{x})$在 $\boldsymbol{x}$ 处是闭的，即若 $\boldsymbol{x}^k \in R^n (k = 1, 2, \cdots)$，$\boldsymbol{x}^k \to \boldsymbol{x}$，$\boldsymbol{z}^k \in \partial^0 f(\boldsymbol{x}^k)$，$\boldsymbol{z}^k \to \boldsymbol{z} (k \to \infty)$，则 $\boldsymbol{z} \in \partial^0 \boldsymbol{f}(\boldsymbol{x})$.

(3) 对任意的 $\varepsilon > 0$，存在 $\delta > 0$ 使得

$$\partial^0 \boldsymbol{f}(\boldsymbol{y}) \subset \partial^0 \boldsymbol{f}(\boldsymbol{x}) + \varepsilon B_{m\times n} \quad \forall \boldsymbol{y} \in \boldsymbol{x} + \delta B.$$

(4) 若 $f_i (i = 1, \cdots, m)$ 在 $\boldsymbol{x}$ 处是 γ_i 秩 Lipschitz 的，则 $\boldsymbol{f}$ 在 $\boldsymbol{x}$ 处是 $\gamma = \| (\gamma_1, \cdots, \gamma_m)^T \|$ 秩 Lipschitz 的，并且 $\partial^0 \boldsymbol{f}(\boldsymbol{x}) \subset \gamma \mathrm{cl} B_{m\times n}$.

(5) $\partial^0 \boldsymbol{f}(\boldsymbol{x}) \subset \partial^0 f_1(\boldsymbol{x}) \times \partial^0 f_2(\boldsymbol{x}) \times \cdots \times \partial^0 f_m(\boldsymbol{x})$，表达式右端意即矩阵的第 i 行的元素属于 $\partial^0 f_i(\boldsymbol{x})$. 当 $m = 1$ 时，$\partial^0 \boldsymbol{f}(\boldsymbol{x}) = \partial^0 f_1(\boldsymbol{x})$.

证明 (1)～(5)都可由定理 4.5.1 推得. □

设 $S \subset R^n$ 是测度为 0 的集合. 在定义 4.6.1 中选取点列 $\{\boldsymbol{x}^k\} \not\subset S$，可得到另一个广义 Jacobi 矩阵：

$$\partial_S^0 \boldsymbol{f}(\boldsymbol{x}) = \mathrm{co}\left\{ \lim_{k\to\infty} J\boldsymbol{f}(\boldsymbol{x}^k) \,|\, \boldsymbol{x}^k \to \boldsymbol{x},\ \boldsymbol{x}^k \notin S \cap \Omega_f \right\}.$$

事实上，可以证明 $\partial_S^0 \boldsymbol{f}(\boldsymbol{x})$与 $\partial^0 \boldsymbol{f}(\boldsymbol{x})$是一致的.

定理 4.6.2 设 $\boldsymbol{f}: R^n \to R^m$ 在 $\boldsymbol{x} \in R^n$ 处是 Lipschitz 的，则对任意的 $\boldsymbol{d} \in R^n$ 和任意的 $\boldsymbol{p} \in R^m$，有

$$\partial^0 f(x)d = \partial_S^0 f(x)d,$$

$$\partial^0 f(x)^T p = \partial_S^0 f(x)^T p.$$

证明 由定义显然有 $\partial_S^0 f(x) \subset \partial^0 f(x)$. 对任意的 $u \in R^m$, 作 $\partial^0 f(x)d$ 的支撑函数 σ_1 和 $\partial_S^0 f(x)d$ 的支撑函数 σ_2:

$$\sigma_1(u) = \lim\sup\{\langle u, Jf(y)d\rangle \mid y \to x,\ y \notin \Omega_f\}$$

$$= \lim\sup\{\langle Jf(y)^T u, d\rangle \mid y \to x,\ y \notin \Omega_f\},$$

$$\sigma_2(u) = \lim\sup\{\langle u, Jf(y)d\rangle \mid y \to x,\ y \notin S \cap \Omega_f\}$$

$$= \lim\sup\{\langle Jf(y)^T u, d\rangle \mid y \to x,\ y \notin S \cap \Omega_f\}.$$

再作函数 $g: R^n \to R$, $g(y) = \langle u, f(y)\rangle$, 注意 ∇g 存在,并当 $y \notin \Omega_f$ 时有 $\nabla g(y)^T = Jf(y)^T u$. 由此,根据定理 4.5.1 得

$$\sigma_1(u) = \lim\sup\{\nabla g(y)^T d \mid y \to x,\ y \notin \Omega_f\} = g^0(x; d)$$

$$= \lim\sup\{\nabla g(y)^T d \mid y \to x,\ y \notin S \cap \Omega_f\} = \sigma_2(u).$$

由定理 4.2.11,即知定理的结论成立. ▯

以下是中值定理和链式法则.

定理 4.6.3 设 $f: R^n \to R^m$ 是开凸集 $S \in R^n$ 上的 Lipschitz 向量函数, $x, y \in S$, 则

$$f(y) - f(x) \in \mathrm{co}\partial^0 f([x, y])(y - x).$$

证明 只要证明存在 $z \in \partial^0 f(\xi)$, $\xi \in [x, y]$, 使得 $f(y) - f(x) \in \mathrm{co}(z(y - x))$. 当固定 $x, y \notin \Omega_f$, 线段 $[x, y]$ 与 Ω_f 的交集的测度为 0,我们有

$$f(y) - f(x) = \int_0^1 Jf(x + (t(y - x))(y - x)\mathrm{d}t,$$

因此, $f(y) - f(x) \in \mathrm{co}\partial^0 f([x, y])(y - x)$. ▯

定理 4.6.4 设 $h: R^n \to R^m$ 在 $x \in R^n$ 处是 Lipschitz 的, $g: R^m \to R$ 在 $h(x)$ 处是Lipschitz的, $f = g \circ h$, 则 f 在 x 处是

Lipschitz的，并且有

$$\partial^0 f(x) \subset \mathrm{co}\{\partial^0 g(h(x)) \circ \partial^0 h(x)\}. \qquad (4.6.1)$$

再若 g 在 x 处是严格可微的，则(4.6.1)成为等式.

证明 容易证明 f 在 x 附近是 Lipschitz 的. 设对任意的 $\epsilon > 0$, $y \in x + \epsilon B$, 有

$$f^0(x; d) \leqslant \frac{f(y+td) - f(y)}{t} + \epsilon,$$

根据定理 4.3.5，我们有

$$f(y+td) - f(y) = \zeta^T(h(y+td) - h(y)),$$

其中 $\zeta \in \partial^0 g(\xi)$, $\xi \in [h(y), h(y+td)]$. 再由定理 4.6.2 可知有 $p \in \mathrm{co}\partial^0 h[y, y+td]d$,

$$h(y+td) - h(y) = tp.$$

由此，存在 $z \in \partial^0 h[y, y+td]$ 使得 $\zeta^T p = \zeta^T(zd)$, 从而

$$f^0(x; d) \leqslant \zeta^T(zd) + \epsilon, \ \zeta^T \in \partial^0 g(\xi), \ \xi \in [h(y), h(y+td)],$$

$$z \in \partial h[y, y+td], \ y \in x + \epsilon B. \qquad (4.6.2)$$

可以选择点列 $\{y^k\} \subset R^n$, $y^k \to x$, $\{t_k\} \in (0, \infty)$, $t_k \to 0^+$ $(k \to \infty)$ 使得(4.6.2)成立. 显然，对应地有 $\xi^k \to h(x)$, $\zeta_k \to \zeta_0 \in \partial^0 g(h(x))$, $z_k \to z_0 \in \partial^0 h(x)$, 由(4.6.2)得到

$$f^0(x; d) \leqslant \zeta_0^T z_0 d + \epsilon.$$

由 ϵ 的任意性，有 $f^0(x; d) \leqslant \zeta_0^T z_0 d$, 再由定理 4.2.11 得知(4.6.1)成立.

现设 g 在 $h(x)$ 处是严格可微的, $D_S g(h(x)) = \zeta$, 对任给的 $\epsilon > 0$, 使得 h 在 $x + \epsilon B$ 上是 γ 秩 Lipschitz 的. 选择 $\delta \in (0, \epsilon)$, 则对任意的 $y \in x + \delta B$, h 在 y 处是 Lipschitz 的, g 在 $h(y)$ 处是 Lipschitz 的. 由定理 4.6.1 的(3)选取充分小的 δ，可使

$$\partial^0 g\{h(y)\} \subset \zeta + \frac{\varepsilon}{\gamma} B. \tag{4.6.3}$$

注意 $\partial^0 g(h(x)) = \{\zeta\}$，对任何向量 $d \in R^n$，当 $y \notin \Omega_f \cup \Omega_h$ 时，有

$$\nabla f(y)^T d = \lim_{t \to 0^+} \frac{g\{h(y+td)\} - g\{h(y)\}}{t}.$$

于是，由定理 4.6.3 可知，存在 $\zeta^t \in \partial^0 g(\xi^t)$ 和 $\xi^t \in [h(y), h(y+td)]$，使得

$$\nabla f(y)^T d = \lim_{t \to 0^+} (\zeta^t)^T \frac{h(y+td) - h(y)}{t}.$$

取 $\{\zeta^t\}$ 的收敛的子列 $\zeta^t \to \zeta$，因为 h 在 y 处可微，$\partial^0 g\{h(y)\} \circ Jh(y)$ 是凸的，从上式得

$$\nabla f(y)^T d \in \partial^0 g\{h(y)\} \circ Jh(y) d. \tag{4.6.4}$$

由(4.6.3)和(4.6.4)，对任意的 $y \in x + \delta B$，$y \notin \Omega_f \cup \Omega_h$，按定理 4.6.1 的(1)，我们有

$$\begin{aligned}\nabla f(y) &\in \partial^0 g\{h(y)\} \circ Jh(y) \\ &\subset \left\{\zeta + \frac{\varepsilon}{\gamma} B\right\} \circ J(h) \subset \zeta \circ Jh(y) + \varepsilon B.\end{aligned} \tag{4.6.5}$$

设对任意的 $d \in R^n$，有

$$\begin{aligned} q &= \max \zeta^T \partial^0 h(x) d \\ &= \lim\sup\{\zeta^T Jh(y) d \mid y \to x,\ y \notin \Omega_h\} \\ &\geqslant \lim\sup\{\zeta^T Jh(y) d \mid y \to x,\ y \notin \Omega_f \cup \Omega_h\}.\end{aligned}$$

从(4.6.5)和推论 4.5.2，我们有

$$q \geqslant \lim\sup\{\nabla f(y)^T d \mid y \to x,\ y \in \Omega_f \cup \Omega_h\},$$

即

$$q \geqslant f^0(x;\ d).$$

再由(4 6.5)得 $q \leqslant f^0(x;\ d)$，故(4.6.1)成为等式. ▯

推论 4.6.5　设 $h:\ R^n \to R^m$ 在 $x \in R^n$ 处是 Lipschitz 的，$g:\ R^m \to R$ 在 $h(x)$处是 Lipschitz 的，则对任意的 $d \in R^n$ 有

$$\partial^0(g \circ h)(x)d \subset \mathrm{co}\{\partial^0 g(h(x)) \circ \partial^0 h(x)d\}. \quad (4.6.6)$$

再若 g 在 $h(x)$处是连续可微的，则(4.6.6)成为等式.

证明　设对任意的 $z,\ y \in R^n$，$g_1(y) = \langle z,\ y\rangle$，记 $f_1 = g \circ h$，$f = g_1 \circ f_1$，应用定理 4.6.4 得

$$\begin{aligned} z^T\partial^0(f_1(x))d &= z^T\partial^0(g \circ h)(x)d \\ &= \partial^0(z^T[g \circ h])(x)d \\ &\subset \mathrm{co}\{\partial^0(z^T g)(h(x) \circ \partial^0 h(x)d)\}. \end{aligned}$$

再根据定理 4.6.4，上式为

$$\begin{aligned} z^T\partial^0 g(h(x))d &\subset \mathrm{co}\{z^T\partial^0 g(h(x)) \circ \partial^0 h(x)d\} \\ &= z^T\mathrm{co}\{\partial^0 g(h(x)) \circ \partial^0 h(x)d\}. \end{aligned}$$

由 z 的任意性，故(4.6.6)成立.

若 g 在 $h(x)$处是连续可微的，则显然有

$$\partial^0(g \circ h)(x) \supset \partial^0 g(h(x)) \circ \partial^0 h(x).$$

因为 $\Omega_{g \cdot h} \subset \Omega_h$，再由定理 4.6.4 得知(4.6.6)成为等式. ▯

第 5 章　函数的切导数和切微分

在第 3 章和第 4 章，我们知道切锥与次梯度有着紧密的联系．因为通过与切锥对应的法锥，可以相应地建立函数的次梯度和次微分．我们还看到，对于函数而言，它的上图象的切锥与方向导数的上图象是一致的．因此，借助切锥也是引进方向导数的一个重要途径．

本章将再介绍几种切锥以及它们的有关性质．由此，引进一般函数的方向导数、次梯度和次微分概念，并得到一些相应的结果．

§5.1　凸集的切锥

在§3.2 中的定义 3.2.2，我们曾利用指示函数的次微分，引进了线性拓扑空间中凸集在一点处的法锥和切锥的概念．现在进而给出在 Banach 空间 $\mathscr{B}$ 中，凸集 $S\subset\mathscr{B}$ 在点 $\boldsymbol{x}\in S$ 处的切锥 $T_S(\boldsymbol{x})$ 和法锥 $N_S(\boldsymbol{x})$ 的等价形式，并且讨论它们的有关性质．

设 $\mathscr{B}^*$ 是 $\mathscr{B}$ 的对偶空间．

定理 5.1.1　设 $S\subset\mathscr{B}$ 是非空凸集，$\boldsymbol{x}\in S$，则

$$T_S(\boldsymbol{x})=\mathrm{cl}\Big(\bigcup_{\lambda>0}\lambda(S-\boldsymbol{x})\Big), \tag{5.1.1}$$

$$N_S(\boldsymbol{x})=\big\{\boldsymbol{x}^*\in\mathscr{B}^*\mid\langle\boldsymbol{x}^*,\boldsymbol{d}\rangle\leqslant 0\quad\forall\boldsymbol{d}\in T_S(\boldsymbol{x})\big\}. \tag{5.1.2}$$

证明　记(5.1.1)和(5.1.2)右端的集合分别为 T 和 N．先证

(5.1.1). 设 $x^* \in T^-$，由定义 1.4.6 的(3)，则对任意的 $d \in T$ 有 $\langle x^*, d\rangle \leqslant 0$，从而对任意的 $y \in S$ 有 $\langle x^*, y - x\rangle \leqslant 0$. 据此，按定义 3.2.2 中的(1)知 $x^* \in N_S(x)$，于是 $T^- \subset N_S(x)$. 由注 3.2.4、注 1.4.3 和注 1.4.4，则得 $T_S(x) = N_S(x)^- \subset (T^-)^- = T$. 反之，设 $d \in T$，则存在 $\lambda_k > 0$, $y^k \in S(k = 1, 2, \cdots)$，使得 $d = \lim\lambda_k(y^k - x)$. 任取 $x^* \in N_S(x)$，由定义 3.2.2 的(1)，则对任意的 k，有 $\langle x^*, y^k - x\rangle \leqslant 0$，从而 $\langle x^*, \lambda_k(y^k - x)\rangle \leqslant 0$. 令 $k \to \infty$，得 $\langle x^*, d\rangle \leqslant 0$. 于是，按定义 1.4.6 的(3)知 $d \in N_S(x)^-$. 因此，再由注 3.2.4 得到 $T \subset N_S(x)^- = T_S(x)$.

为证(5.1.2)，设 $x^* \in N$，由(5.1.1)可知对任意的 $y \in S$，有 $\langle x^*, y - x\rangle \leqslant 0$. 故按定义 3.2.2 的(1)，得 $x^* \in N_S(x)$. 反之，设 $x^* \in N_S(x)$，根据定义 3.2.2 的(1)有

$$\langle x^*, y - x\rangle \leqslant 0 \quad \forall y \in S.$$

由此，设 $d = \lim\limits_{k\to\infty}\lambda_k(y^k - x) \in T_S(x)$，其中 $\lambda_k > 0$, $y^k \in S(k = 1, 2, \cdots)$，则对任意的 k，有 $\langle x^*, y^k - x\rangle \leqslant 0$，从而 $\langle x^*, d\rangle \leqslant 0$. 于是，得到 $x^* \in N$. □

注 5.1.1　(5.1.1)和(5.1.2)分别是 Banach 空间中凸集的切锥和法锥的等价形式. 设 $B \subset \mathscr{B}$ 是开单位球，由(5.1.1)容易知道 $d \in T_S(x)$ 等价于：对任意的 $\varepsilon > 0$，存在 $p \in d + \varepsilon B$ 和 $t > 0$ 使得 $x + tp \in S$. 它也等价于：存在序列 $\{d^k\} \subset S$, $d^k \to d$ 和 $\{t_k\} \subset (0, +\infty)$, $t_k \to 0(k \to \infty)$，使得对任意的 $k = 1, 2, \cdots$，有 $x + t_k d^k \in S$. 此外，它也等价于下式成立：

$$\liminf_{t_k\to 0} \frac{1}{t_k} d_S(x + t_k d^k) = 0,$$

其中 $d^k \to d(k \to \infty)$.

定理 5.1.2　设 $S \subset \mathscr{B}$ 是非空凸集，$x \in S$.

(1) $T_S(x) = T_{\mathrm{cl}S}(x)$.

(2) $T_S(x) = \mathscr{B}$ $(\forall x \in \mathrm{int}S)$.

(3) $S \subset x + T_S(x)$.

证明 由(5.1.1)直接可得证. ▯

定理 5.1.3 设 $S \subset \mathscr{B}$ 是闭凸集, $x \in S$, 则 $N_S(x)$是闭集.

证明 设点列 $\{x_k\} \subset T_S(x)$, $x_k \to x$, $\{x_k^*\} \in N_S(x_k)$, $x_k^* \to x^*$ $(k \to \infty)$. 根据定义 3.2.2 的(2), 对任意的 $d \in S$, 有 $\langle x_k^*, d\rangle \leqslant \langle x_k^*, x_k\rangle$. 令 $k \to \infty$, 得

$$\langle x^*, d\rangle \leqslant \langle x^*, x\rangle \quad \forall d \in S,$$

故 $x^* \in N_S(x)$. ▯

定理 5.1.4 设 $S \subset \mathscr{B}$ 是凸集. 若 $\mathrm{int}S \neq \varnothing$, 则

$$\mathrm{int}T_S(x) = \bigcup_{\lambda>0}\lambda(\mathrm{int}S - x) \quad \forall x \in S, \tag{5.1.3}$$

并且映射: $x \to \mathrm{int}T_S(x)$ 的图象是开集.

证明 显然 $\bigcup_{\lambda>0}\lambda(\mathrm{int}S - x)$ 是开集, 并且 $\bigcup_{\lambda>0}\lambda(\mathrm{int}S - x) \subset \mathrm{int}T_S(x)$. 另外, 因为 $\mathrm{int}T_S(x) = \mathrm{int}C_S(x)$ (其中 $C_S(x) = \bigcup_{\lambda>0}\lambda(S - x)$), 设 $d \in \mathrm{int}C_S(x)$, 可取 $\delta > 0$ 使 $d + \delta B \subset C_S(x)$. 若 $d \notin S - x$. 设 $x^0 \in \mathrm{int}S$, $d^0 = x^0 - x$, 有 $d - \frac{\delta d^0}{\|d^0\|} \in C_S(x)$, 则存在 $\lambda > 0$ 使得

$$x + \frac{1}{\lambda}\left(d - \frac{\delta d^0}{\|d^0\|}\right) \in S.$$

令 $\alpha = \frac{\delta}{\delta + \lambda\|d^0\|}$, 我们有

$$\lambda x + (1-\alpha)d = \alpha\lambda x^0 + (1-\alpha)\left(\lambda x + \left(d - \frac{\delta d^0}{\|d^0\|}\right)\right) \in S.$$

从上式可得 $d \in \frac{\lambda}{(1-\alpha)}(\mathrm{int}S - x)$, 因而(5.1.3)成立.

设 $d^0 \in \mathrm{int}T_S(x_0)$, 则存在某个 $\lambda_0 > 0$, 使 $d^0 \in \lambda_0(\mathrm{int}S - x^0)$, 因此存在 $\varepsilon > 0$ 使得

$$x^0 + \lambda_0 d^0 + \varepsilon B = x^0 + \lambda_0\left(d^0 + \frac{\varepsilon}{\lambda_0}B\right) \subset \mathrm{int}S.$$

取 $\boldsymbol{x} \in \boldsymbol{x}^0 + \frac{\varepsilon}{2}B$ 和 $\boldsymbol{d} \in \boldsymbol{d}^0 + \frac{\varepsilon}{2\lambda_0}B$，由上式得

$$\boldsymbol{x} + \lambda_0 \boldsymbol{d} \in \boldsymbol{x}^0 + \lambda_0 \boldsymbol{d}^0 + \varepsilon B \subset \mathrm{int}S,$$

从而 $\boldsymbol{d} \in \mathrm{int}T_S(\boldsymbol{x})$. 由此，$\boldsymbol{x} \to \mathrm{int}T_S(\boldsymbol{x})$的图象是开集. ▯

定理 5.1.5 设 $\mathscr{H}$ 是 Hilbert 空间，$B \subset \mathscr{H}$ 是闭单位球，$\boldsymbol{x} \in B$.

(1) 若 $\boldsymbol{x} \in \mathrm{int}B$，则 $T_B(\boldsymbol{x}) = \mathscr{H}$.

(2) 若 $\boldsymbol{x} \in \mathrm{int}B$，则 $N_B(\boldsymbol{x}) = \{\boldsymbol{0}\}$.

证明 由定理 5.1.1 可得证. ▯

设 S^- 表示集合 $S \subset \mathscr{B}$ 的负对偶集(或称极化集)，即 $S^- = \{\boldsymbol{x}^* \in \mathscr{B}^* \mid \langle \boldsymbol{x}^*, \boldsymbol{x}\rangle \leqslant 0 \quad \forall \boldsymbol{x} \in S\}$ (若 $S = K$ 是凸锥，由定义 1.4.6 的(3)，K^- 是 K 的负对偶锥). $S^\perp$表示集合 S 的垂直集，即

$$S^\perp = \{\boldsymbol{x}^* \in \mathscr{B}^* \mid \langle \boldsymbol{x}^*, \boldsymbol{x}\rangle = 0 \quad \forall \boldsymbol{x} \in S\}.$$

定理 5.1.6 设 $K \subset \mathscr{B}$ 是闭凸锥，$\boldsymbol{x} \in \mathscr{B}$，则 $N_K(\boldsymbol{x}) = K^- \cap \{\boldsymbol{x}\}^\perp$，即 $\boldsymbol{d} \in T_K(\boldsymbol{x})$ 当且仅当对任意的 $\boldsymbol{x}^* \in K^-$，满足 $\langle \boldsymbol{x}^*, \boldsymbol{x}\rangle = 0$ 使得 $\langle \boldsymbol{x}^*, \boldsymbol{d}\rangle \leqslant 0$. 再若 $S \subset \mathscr{B}$ 是闭线性子空间，则 $T_K(\boldsymbol{x}) = K$，$N_K(\boldsymbol{x}) = K^\perp$.

证明 显然有 $K^- \cap \{\boldsymbol{x}\}^\perp \subset N_K(\boldsymbol{x})$. 反之，若 $\boldsymbol{x}^* \in N_K(\boldsymbol{x})$，根据定义 3.2.2 的(1)有 $\langle \boldsymbol{x}^*, \boldsymbol{x}\rangle = \max\limits_{\boldsymbol{d} \in K} \langle \boldsymbol{x}^*, \boldsymbol{d}\rangle$，因为 K 是锥，所以有 $\langle \boldsymbol{x}^*, \boldsymbol{x}\rangle = 0$ 和 $\langle \boldsymbol{x}^*, \boldsymbol{d}\rangle \leqslant 0 (\forall \boldsymbol{d} \in K)$. ▯

设 $\mathscr{D}$ 是 Banach 空间，$\varphi: \mathscr{B} \to \mathscr{D}$ 是线性映射，$\mathrm{Ker}\varphi = \{\boldsymbol{x} \in \mathscr{B} \mid \varphi(\boldsymbol{x}) = \boldsymbol{0}\}$ 是 φ 的核.

定理 5.1.7 设 $\varphi: \mathscr{B} \to \mathscr{D}$ 是线性映射，$S = \varphi^{-1}(\boldsymbol{y}) \subset \mathscr{B}$ 是仿射子空间. 若 $\varphi(\boldsymbol{x}) = \boldsymbol{y}$，则 $T_S(\boldsymbol{x}) = \mathrm{Ker}\varphi$.

证明 若 $\boldsymbol{d} \in \mathrm{Ker}\varphi$，则 $\boldsymbol{d} + \boldsymbol{x} \in \varphi^{-1}(\boldsymbol{y}) = S$，因此 $\boldsymbol{d} \in T_S(\boldsymbol{x})$. 反之，设 $\boldsymbol{d}^k = \lambda_k(\boldsymbol{x}^k - \boldsymbol{x})$，$\boldsymbol{x}^k \in S$，$\boldsymbol{d}^k \to \boldsymbol{d} \in T_S(\boldsymbol{x}) (k \to \infty)$，则 $\varphi(\boldsymbol{d}^k) = \lambda_k \varphi(\boldsymbol{x}^k - \boldsymbol{x}) = \boldsymbol{0}$. 由此，得 $\boldsymbol{d}^k \in \mathrm{Ker}\varphi$，从而 $\boldsymbol{d} \in \mathrm{Ker}\varphi$. ▯

定理 5.1.8 设集合

$$S = \left\{ \boldsymbol{x} \in R_+^n \,\middle|\, \boldsymbol{x} = (x_1, \cdots, x_n)^T,\ \sum_{i=1}^n x_i = 1 \right\},$$

$$I(\boldsymbol{x}) = \{ i = 1, \cdots, n \mid x_i = 0 \},$$

$$\boldsymbol{d} = (d_1, \cdots, d_n)^T.$$

(1) $\boldsymbol{d} \in T_{R_+^n}(\boldsymbol{x})$ 当且仅当对任何 $i \in I(\boldsymbol{x})$ 有 $d_i \geqslant 0$.

(2) $\boldsymbol{d} \in T_S(\boldsymbol{x})$ 当且仅当对任何 $i \in I(\boldsymbol{x})$ 有 $d_i \geqslant 0$，并且 $\sum_{i=1}^n d_i = 0$.

证明 (1) 由定理 5.1.6 得，当 $\boldsymbol{x}^* = (x_1^*, \cdots, x_n^*)^T$ 时，有

$$\sum_{i=1}^n x_i^* x_i = \sum_{i \notin I(x)} x_i^* x_i = 0,$$

从而当 $i \notin I(\boldsymbol{x})$ 时有 $x_i^* = 0$. 据此，即得 $\boldsymbol{d} \in T_{R_+^n}(\boldsymbol{x})$ 当且仅当 $\sum_{i \notin I(x)} x_i^* d_i \leqslant 0\ (\forall x_i^* \in R^n)$，即对任何 $i \in I(\boldsymbol{x})$，有 $d_i \geqslant 0$.

(2) 对任何 $i \in I(\boldsymbol{x})$，设 $\boldsymbol{d} = (d_1, \cdots, d_n)^T$ 满足 $d_i \geqslant 0$，并且 $\sum_{i=1}^n d_i = 0$. 若 $\boldsymbol{d} = \boldsymbol{0}$，则 $\boldsymbol{d} \in T_S(\boldsymbol{x})$. 若 $\boldsymbol{d} \neq \boldsymbol{0}$，设 $\lambda = \lim\limits_{\substack{i \notin I(x) \\ d_i \neq 0}} \dfrac{x_i}{|d_i|} > 0$，我们有

$$x_i + \lambda d_i = \lambda d_i \geqslant 0,\ i \in I(\boldsymbol{x}),$$

$$x_i + \lambda d_i \geqslant x_i - \lambda |d_i| \geqslant 0,\ i \notin I(\boldsymbol{x}),$$

$$\sum_{i=1}^n (x_i + \lambda d_i) = 1 + 0 = 1.$$

因此，得到 $\boldsymbol{x} + \lambda \boldsymbol{d} \in S$，故 $\boldsymbol{d} \in T_S(\boldsymbol{x})$. 反之，对任意的 $\boldsymbol{y} = (y_1, \cdots, y_n)^T \in S$ 和 $\lambda > 0$，令 $\boldsymbol{d} = \lambda(\boldsymbol{y} - \boldsymbol{x})$，则有

$$d_i = \lambda(y_i - x_i) = \lambda y_i \geqslant 0 \quad \forall i \in I(\boldsymbol{x})$$

和 $\sum_{i=1}^{n} d_i = 0$. 因此有 $\bigcup_{\lambda>0} \frac{1}{\lambda}(S - \boldsymbol{x}) \subset \left\{ \boldsymbol{d} \in T_{R_+^n}(\boldsymbol{x}) \,\middle|\, \sum_{i=1}^{n} d_i = 0 \right\}$.
注意后者是闭的，所以 $T_S(\boldsymbol{x}) = \left\{ \boldsymbol{d} \in T_{R_+^n}(\boldsymbol{x}) \,\middle|\, \sum_{i=1}^{n} d_i = 0 \right\}$. □

定理 5.1.9　设 $S_1, S_2 \subset \mathscr{B}$, $S_i \subset \mathscr{B}(i \in I, I$ 是指标集$)$是凸集，$\boldsymbol{x} \in \mathscr{B}$.

(1) 若 $\boldsymbol{x} \subset S_1 \subset S_2$，则 $T_{S_1}(\boldsymbol{x}) \subset T_{S_2}(\boldsymbol{x})$ 和 $N_{S_2}(\boldsymbol{x}) \subset N_{S_1}(\boldsymbol{x})$.

(2) 若 $S = \bigcap_{i \in I(\boldsymbol{x})} S_i$ 和 $I(\boldsymbol{x}) = \{ i \mid \boldsymbol{x} \notin \mathrm{int} S_i \}$，则 $T_S(\boldsymbol{x}) \subset \bigcap_{i \in I(\boldsymbol{x})} T_{S_i}(\boldsymbol{x})$.

证明　由定理 5.1.1 可直接推得结论成立. □

定理 5.1.10　设 $S_i \subset \mathscr{B}(i = 1, \cdots, m)$ 是凸集，$S = \prod_{i=1}^{m} S_i$，$\boldsymbol{x}^i \in S_i(i = 1, \cdots, m)$，$\boldsymbol{x} = (\boldsymbol{x}^1, \cdots, \boldsymbol{x}^m)$，则

$$T_S(\boldsymbol{x}) = \prod_{i=1}^{m} T_{S_i}(\boldsymbol{x}^i), \quad N_S(\boldsymbol{x}) = \prod_{i=1}^{m} N_{S_i}(\boldsymbol{x}^i).$$

证明　显然有 $T_S(\boldsymbol{x}) \subset \prod_{i=1}^{m} T_{S_i}(\boldsymbol{x}^i)$. 反之，设 $\boldsymbol{d}^i \in T_{S_i}(\boldsymbol{x}^i)(i = 1, \cdots, m)$，则存在 $\boldsymbol{d}_k^i \to \boldsymbol{d}^i (k \to \infty)$和 $\lambda_i^k > 0$，使得 $\boldsymbol{x}^i + \lambda_i^k \boldsymbol{d}_k^i \in S_i(i = 1, \cdots, m)$. 令 $\lambda = \min\{\lambda_k^1, \cdots, \lambda_k^m\}$，由于 S_i 是凸的，我们有

$$\boldsymbol{x}^i + \lambda_i \boldsymbol{d}_i^k = \left(1 - \frac{\lambda_i}{\lambda_i^k}\right)\boldsymbol{x}^i + \frac{\lambda_i}{\lambda_i^k}(\boldsymbol{x}^i + \lambda_i^k \boldsymbol{d}_i^k) \in S_i,$$

$$i = 1, \cdots, m.$$

据此，得到 $\boldsymbol{x} + \lambda \boldsymbol{d}^i \in S$，其中 $\boldsymbol{d}^i = (\boldsymbol{d}_1^i, \cdots, \boldsymbol{d}_m^i)$，因此 $\boldsymbol{d} \in T_S(\boldsymbol{x})$.

同理，由(5.1.2)可以推出第二个结论成立. □

定理 5.1.11　设 $\varphi: \mathscr{B} \to \mathscr{D}$ 是线性映射，$S \subset \mathscr{B}$ 是凸集，$\boldsymbol{x} \in S$.

(1) $T_{\varphi(S)}(\varphi(\boldsymbol{x})) = \mathrm{cl}\varphi(T_S(\boldsymbol{x}))$.

(2) $N_{\varphi(S)}(\boldsymbol{x}) = \varphi^{*-1}(N_S(\boldsymbol{x}))$ (φ^* 是 φ的对偶算子).

证明 因为$\langle x^*, \varphi(x)\rangle = \max_{d\in S}\langle x^*, \varphi(d)\rangle = \max_{d\in S}\langle \varphi^*(x^*), d\rangle = \langle \varphi^*(x^*), x\rangle$，所以(2)成立. 由(2)可以推出(1)成立. □

推论 5.1.12 设$S_1, S_2 \subset \mathscr{B}$是闭凸集，$x^1 \in S_1$，$x^2 \in S_2$. 则

$$T_{S_1+S_2}(x^1 + x^2) = \mathrm{cl}(T_{S_1}(x^1) + T_{S_2}(x^2)),$$

$$N_{S_1+S_2}(x^1 + x^2) = N_{S_1}(x^1) \cap N_{S_2}(x^2).$$

证明 由定理 5.1.11 可以推得结论成立. □

定理 5.1.13 设$S \subset \mathscr{B}$和$M \subset \mathscr{D}$是闭凸集，$\varphi: \mathscr{B} \to \mathscr{D}$是线性映射，$V = \{x \in \mathscr{B} \mid \varphi(x) \in M\} = S \cap \varphi^{-1}(M)$. 又设$0 \in \varphi(S) - M$（即$V \neq \varnothing$），$x \in V$，以及

$$T_V(x) \subset T_S(x) \cap \varphi^{-1}\big(T_M(\varphi(x))\big),$$

$$N_S(x) + \varphi^*\big(N_M(\varphi(x))\big) \subset N_V(x). \qquad (5.1.4)$$

若$0 \in \mathrm{int}(\varphi(S) - M)$，$x \in V$，则

(1) $T_V(x) = T_S(x) \cap \varphi^{-1}\big(T_M(\varphi(x))\big)$；

(2) $N_V(x) = N_S(x) + \varphi^*\big(N_M(\varphi(x))\big)$.

证明 (1) 设$d \in T_S(x) \cap \varphi^{-1}\big(T_M(\varphi(x))\big)$，则存在点列$\{d^k\} \subset \mathscr{B}$，$d^k \to d$，$\{y^k\} \subset \mathscr{D}$，$y^k \to \varphi(d)$，以及$\{\lambda_k^1\}$和$\{\lambda_k^2\} \subset (0, +\infty)$，$\lambda_k^1 \to 0$，$\lambda_k^2 \to 0 (k \to \infty)$，使得对所有的$k$，有

$$x + \lambda_k^1 d^k \in S, \quad \varphi(x) + \lambda_k^2 y^k \in M.$$

令$\lambda_k = \min(\lambda_k^1, \lambda_k^2)$，因为$S$和$M$是凸的，可以推得

$$x + \lambda_k d^k \in S, \quad \varphi(x) + \lambda_k y^k \in M, \ k = 1, 2, \cdots.$$

若$y^k = \varphi(d^k)(k = 1, 2, \cdots)$中有无穷多个成立，则有$d \in T_V(x)$. 否则，作集值映射$\psi: S \to 2^M$，$x \mapsto \psi(x)$，

$$\psi(x) = \varphi(x) - M.$$

根据已知$0 \in \mathrm{int}\psi(S) = \mathrm{int}(\varphi(S) - M)$，由文献[16]的推论

3.3.4，存在 $\delta>0$ 使得对任意的 $y'\in\mathscr{B}$ 和 $x'\in S$，有

$$d_{\varphi^{-1}(y')}(x')\leqslant\frac{1}{\delta}d_{\varphi(y')}(x')\left(\|x'-x\|+1\right).$$

取 $y'=\boldsymbol{0}$，$x'=x+\lambda_k d$，则有 $\|x'-x\|=\lambda_k\|d\|$ 和

$$\begin{aligned}d_{\varphi(x+\lambda_k d)}(\boldsymbol{0})&=d_M\left(\varphi(x)+\lambda_k\varphi(d)\right)\\&\leqslant d_M\left(\varphi(x)+\lambda_k y^k\right)+\lambda_k\|\varphi(d)-y^k\|\\&=\lambda_k\|\varphi(d)-y^k\|.\end{aligned}$$

于是，我们有

$$\begin{aligned}\frac{1}{\lambda_k}d_V(x+\lambda_k d^k)&\leqslant\frac{1}{\delta\lambda_k}d_{\varphi(x+\lambda_k d)}(\boldsymbol{0})\left(\|x'-x\|+1\right)\\&\leqslant\frac{1}{\delta}\|\varphi(x)-y^k\|\left(\lambda_k\|d\|+1\right).\end{aligned}$$

因为 $y^k\to\varphi(d)\,(k\to\infty)$，所以从上式知

$$\liminf_{\lambda_k\to\infty}\frac{1}{\lambda_k}d_V(x+\lambda_k d^k)=0.$$

由注 5.1.1，即得 $d\in T_V(x)$.

(2) 由(1)按(5.1.2)有

$$N_V(x)=\mathrm{cl}\left(N_S(x)+\varphi^*\left(N_M(\varphi(x))\right)\right).$$

现在只要证明 $N_S(x)+\varphi^*\left(N_M(\varphi(x))\right)$ 是闭集即可. 设 $p_k^*\in N_S(x)$，$q_k^*\in N_M(\varphi(x))$，$r_k=p_k^*+\varphi^*(q_k^*)$，$r_k\to r\,(k\to\infty)$，由 $\boldsymbol{0}\in\mathrm{int}(\varphi(S)-M)$ 可知，存在 $\bar{\lambda}>0$，$\bar{d}\in S$ 和 $\bar{y}\in M$，使得对任意的 $y\in\mathscr{D}$ 有 $y=\bar{\lambda}\left(\bar{y}-\varphi(\bar{d})\right)$，因此

$$\begin{aligned}\langle q_k^*,y\rangle&=\bar{\lambda}\left\langle q_k^*,\bar{y}-\varphi(\bar{d})\right\rangle\\&=\bar{\lambda}\left(\langle q_k^*,\bar{y}\rangle-\langle\varphi^*(q_k^*),\bar{d}\rangle\right)\end{aligned}$$

$$= \bar{\lambda}\big(\langle p_k^*, \bar{d}\rangle + \langle q_k^*, \bar{y}\rangle - \langle r_k, \bar{d}\rangle\big)$$
$$\leqslant \bar{\lambda}\big(\langle p_k^*, x\rangle + \langle q_k^*, \varphi(x)\rangle - \langle r_k, \bar{d}\rangle\big)$$
$$= \bar{\lambda}\langle r_k, x - \bar{d}\rangle \leqslant \bar{\lambda}\| r_k \| \cdot \| x - \bar{d} \| < +\infty.$$

根据 Banach-Steinhaus 定理(见文献[10]P. 321 的定理 1),$\{q_k^*\}$ 中存在弱收敛的子列,不妨设 $q_k^* \to q^* \in N_M(\varphi(x))$. 因此,

$$p_k^* = r_k - \varphi^*(q_k^*) \to p^* = r - \varphi^*(q_k^*) \in N_S(x),$$

故得 $r \in N_S(x) + \varphi^*\big(N_M(\varphi(x))\big)$. ▯

由定理 5.1.13,容易推得以下三个推论.

推论 5.1.14 设 $S \subset \mathscr{B}$ 是闭凸集,$\varphi: \mathscr{B} \to \mathscr{D}$ 是线性映射,则对任意的 $y \in \mathrm{int}\varphi(S)$ 和 $x \in S \cap \varphi^{-1}(y)$ 有

$$T_{S\cap\varphi^{-1}(y)}(x) = T_S(x) \cap \mathrm{Ker}\varphi,$$
$$N_{S\cap\varphi^{-1}(y)}(x) = N_S(x) + \mathrm{Im}\varphi^*,$$

其中 $\mathrm{Ker}\varphi$ 是 φ 的核,φ^* 是 φ 的对偶算子,

$$\mathrm{Im}\varphi^* = \{\varphi(x^*) \mid \langle x^*, \varphi(x)\rangle = \langle \varphi^*(x^*), x\rangle, x \in S\}.$$

推论 5.1.15 设 $S \subset \mathscr{B}$ 是闭凸集,$x \in S$,$p_0^* \in S^-$,$V = \{x \in S \mid \langle p_0^*, x\rangle = -1\} \neq \varnothing$,则 $d \in T_V(x)$ 当且仅当对任意的 $p^* \in S^-$ 满足 $\langle p^*, x\rangle = 0$,有 $\langle p_0^*, d\rangle = 0$ 和 $\langle p^*, d\rangle \leqslant 0$.

推论 5.1.16 设 $S_1, S_2 \subset \mathscr{B}$ 是闭凸集.若 $\theta \in \mathrm{int}(S_1 - S_2)$,则

$$T_{S_1\cup S_2}(x) = T_{S_1}(x) \cap T_{S_2}(x) \quad \forall x \in S_1 \cap S_2.$$

再给出如下定理:

定理 5.1.17 设 $S_1, \cdots, S_m \in \mathscr{B}$ 是闭凸集,$S = \bigcap_{i=1}^{m} S_i$. 若存在 $\bar{\delta} > 0$ 使得对任意的 $x^i \in \bar{\delta}B(i = 1, \cdots, m)$ 有 $\bigcap_{i=1}^{m}(S_i - x^i) \neq \varnothing$,则

$$T_S(x) = \bigcap_{i=1}^{m} T_{S_i}(x) \quad \forall x \in S.$$

证明　由推论 5.1.14 和定理 5.1.10 容易推得结论成立. □

§ 5.2　D-切锥和 C-切锥

现在介绍一般集合在一点处的切锥.

设 $\mathscr{B}$ 是 Banach 空间, $\mathscr{B}^*$ 是 $\mathscr{B}$ 的对偶空间, 集合 $S \subset \mathscr{B}$, 点 $x \in S$. $\varepsilon > 0$, 记 $B_S(x, \varepsilon) = S \cap (x + \varepsilon B)$, $x^k \to_S x$ 表示点列 $\{x^k\} \subset S$ 收敛于 $x \in S$.

定义 5.2.1　设集合 $S \subset \mathscr{B}$ 非空, 点 $x^0 \in \mathscr{B}$. 称集合

$$T_S^D(x^0) = \bigcap_{\varepsilon > 0} \bigcap_{\alpha > 0} \bigcup_{0 < \lambda \leqslant \alpha} \left(\frac{1}{\lambda}(S - x^0) + \varepsilon B \right) \qquad (5.2.1)$$

是集合 S 在点 x^0 处的相依切锥, 或 D-切锥.

注 5.2.1　设 $x \in \mathscr{B}$, 由 (5.2.1) 不难得知有以下三个等价关系:

(1) $d \in T_S^D(x)$ 当且仅当对任意的 $\varepsilon > 0$ 和任意的 $\alpha > 0$, 分别存在 $x \in d + \varepsilon B$ 和 $t \in [0, \alpha]$ 使得 $x + td \in S$.

(2) $d \in T_S^D(x)$ 当且仅当存在序列 $\{d^k\} \subset \mathscr{B}$, $d^k \to d$, 以及 $\{t_k\} \subset (0, +\infty)$, $t_k \to 0 (k \to \infty)$, 使得对任何 $k = 1, 2, \cdots$ 有 $x + t_k d^k \in S$.

(3) $d \in T_S^D(x)$ 当且仅当 $\liminf\limits_{t \to 0^+} \dfrac{d_S(x + td)}{t} = 0$.

注 5.2.2　设 $x \in \mathscr{B}$, 由定义 5.2.1 易知 $T_S^D(x)$ 是闭锥. 并且, 当 $x \in \mathrm{int} S$ 时有 $T_S^D(x) = \mathscr{B}$, $T_{\mathscr{B}}^D(x) = \mathscr{B}$ 和 $T_{\varnothing}^D(x) = \varnothing$.

设集合 $S \subset \mathscr{B}$, 点 $x \in S$, $\psi: S \to 2^{\mathscr{B}}$, $x \mapsto \psi(x)$ 是集值映射. 记

$$\liminf_{x' \to x} \psi(x') = \bigcap_{\varepsilon > 0} \bigcup_{\alpha > 0} \bigcap_{x' \in B_S(x, \alpha)} \left(\psi(x') + \varepsilon B \right),$$

它等价于:对任意的 $\varepsilon > 0$, 存在 $\eta > 0$ 使得 $\sup\limits_{x' \in B_S(x, \eta)} d_{\varphi(x')}(d) \leqslant \varepsilon$.

定义 5.2.2 设集合 $S \subset \mathscr{B}$ 非空,点 $x^0 \in \mathscr{B}$.

(1) 集合

$$T_S^C(x^0) = \liminf_{\substack{x \to x^0 \\ \lambda \to 0^+}} \frac{1}{\lambda}(S - x)$$

$$= \bigcap_{\varepsilon > 0} \bigcup_{\alpha, \beta > 0} \bigcap_{\substack{x \in B_S(x^0, \alpha) \\ \lambda \in (0, \beta)}} \left\{ \frac{1}{\lambda}(S - x) + \varepsilon B \right\} \qquad (5.2.2)$$

称为集合 S 在点 x^0 处的 Clarke 切锥,或 C-切锥.

(2) 集合

$$N_S^C(x^0) = \{x^* \in \mathscr{B}^* \mid \langle x^*, d \rangle \leqslant 0, d \in T_S^C(x^0)\} \qquad (5.2.3)$$

称为集合 S 在点 x^0 处的 Clarke 法锥,或 C-法锥.

注 5.2.3 设 $x \in \mathscr{B}$, 由(5.2.2),下面三个等价关系显然成立:

(1) $d \in T_S^C(x)$ 当且仅当对任意的 $\varepsilon > 0$, 存在 $\alpha > 0$ 和 $\beta > 0$, 使对任意的 $x' \in B_S(x, \alpha)$ 和任意的 $\lambda \in [0, \beta]$, 存在 $p \in d + \varepsilon B$ 使得 $x' + \lambda p \in S$.

(2) $d \in T_S^C(x)$ 当且仅当对任何序列 $\{x^k\} \subset S$, $x^k \to x$, $\{t_k\} \in (0, +\infty)$, $t_k \to 0(k \to \infty)$, 存在序列 $\{d^k\} \subset \mathscr{B}$, $d^k \to d$, 使得对任何 $k = 1, 2, \cdots$, 有 $x^k + t_k d^k \in S$.

(3) $d \in T_S^C(x)$ 当且仅当 $\lim\limits_{\substack{x' \to_S x \\ t \to 0^+}} \dfrac{d_S(x' + td)}{t} = 0$.

注 5.2.4 显然,当 $x \in \text{int}S$ 时有 $T_S^C(x) = \mathscr{B}$;当 $x \in \mathscr{B}$ 时有 $T_{\mathscr{B}}^C(x) = \mathscr{B}$ 和 $T_{\varnothing}^C(x) = \varnothing$.

注 5.2.5 当 $S \subset \mathscr{B}$ 非空时,由定理 4.4.7 和注 5.2.3 的(2)可知,C-切锥与 G-切锥是一致的.

定理 5.2.1 设集合 $S \subset \mathscr{B}$ 非空, $x \in \mathscr{B}$, 则 $T_S^C(x)$是闭

凸集.

证明 设 $d^1, d^2 \in T_S^C(x)$，对任何收敛于$(x, 0)$的序列$\{(x^k, \lambda_k)\} \subset S \times (0, +\infty)$，存在序列$\{d_k^1\} \subset \mathscr{B}$收敛于$d^1$，使得对任何$k = 1, 2, \cdots$，有$d_k = x^k + \lambda_k d_k^1 \in S$. 因为$k \to \infty$时$d_k \to x$，由$d^2 \in T_S^C(x)$知存在序列$\{d_k^2\} \subset \mathscr{B}$收敛于$d^2$，使得对任何$k = 1, 2, \cdots$，有

$$d_k + \lambda_k d_k^2 = x^k + \lambda_k(d_k^1 + d_k^2) \in S, k = 1, 2, \cdots.$$

从$d_k^1 + d_k^2 \to d^1 + d^2$知$d^1 + d^2 \in T_S^C(x)$，因此$T_S^C(x)$是闭凸集. ▯

注 5.2.6 这里定义的切锥$T_S^D(x)$和$T_S^C(x)$与第 4 章中引进的 G-切锥有如下关系：

$$T_S^C(x) \subset T_S^D(x) \subset T_S^0(x). \tag{5.2.4}$$

当 S 是凸集时，(5.2.4)的包含关系成为等式.

定理 5.2.2 设集合 $S \subset \mathscr{B}$ 非空，$x \in \mathscr{B}$. 若 S 是凸集，则

$$T_S^C(x) = T_S^D(x) = T_S^0(x) = T_S(x). \tag{5.2.5}$$

证明 由定理 4.4.5 和(5.2.4)，只需证明 $T_S(x) \subset T_S^C(x)$. 设 $d \in T_S(x)$，由(5.1.1)有 $d \in \mathrm{cl} \bigcup_{\lambda>0} \frac{1}{\lambda}(S - x)$. 取任给的 $\varepsilon > 0$，则存在 $y \in S$ 和 $\beta > 0$，使得 $d - \frac{1}{\beta}(y - x) \in (\varepsilon/2)B$. 令 $\alpha = (\varepsilon/2)\beta$，$x' \in B_S(x, \alpha)$，$\lambda \in (0, \beta)$ 和 $p = (y - x')/\beta$，则有

$$x + \lambda p = (1 - \lambda/\beta)x + (\lambda/\beta)y \in S,$$

$$\begin{aligned}\| p - d \| &\leqslant (1/\beta) \| x' - x \| + \| d - (y - x')/\beta \| \\ &\leqslant \alpha/\beta + \varepsilon/2 = \varepsilon.\end{aligned}$$

因此，按定义 5.2.2 的(1)得 $d \in T_S^C(x)$. 定理得证. ▯

通过下面的例子可以看到 $T_S^C(x)$和 $T_S^D(x)$不一定相等.

例 5.2.1 考虑函数 $f: R \to R$，当 $x \leqslant 0$ 时 $f(x) = 0$，当 $x \geqslant$

0 时 $f(x)=x$. 取 $S=\{(x,f(x))\mid x\in R\}$.

若 $x<0$, 则 $T_S^C(x,0)=T_S^D(x,0)=R\times\{0\}$.

若 $x=0$, 则 $T_S^C(0,0)=(0,0)$,

$$T_S^D(0,0)=(-R_+\times\{0\})\cup\{(x',x')\mid x'\geqslant 0\}.$$

若 $x>0$, 则 $T_S^C(x,x)=T_S^D(x,x)=\{(x',x')\mid x'\in R\}$.

定理 5.2.3 设集合 $S\subset R^n$ 非空,则对任意的 $\boldsymbol{x}\in S$, 有

$$T_S^C(\boldsymbol{x})\subset\liminf_{\boldsymbol{x}'\to\boldsymbol{x}} T_S^D(\boldsymbol{x}).$$

证明 由定义 5.2.2,有

$$T_S^C(\boldsymbol{x})=\bigcap_{\varepsilon>0}\bigcup_{\alpha,\beta>0}\bigcap_{\boldsymbol{x}'\in B_S(\boldsymbol{x},\alpha)}\bigcap_{\lambda\in(0,\beta)}\left(\frac{1}{\lambda}(S-\boldsymbol{x}')+\varepsilon B\right).$$

固定 ε 和 α,则

$$\bigcup_{\beta>0}\bigcap_{\boldsymbol{x}'\in B_S(\boldsymbol{x},\alpha)}\bigcap_{\lambda\in(0,\beta)}\left(\frac{1}{\lambda}(S-\boldsymbol{x}')+\varepsilon B\right)$$

$$\subset\bigcap_{\boldsymbol{x}'\in B_S(\boldsymbol{x},\alpha)}\bigcup_{\beta>0}\bigcap_{\lambda\in(0,\beta)}\left(\frac{1}{\lambda}(S-\boldsymbol{x}')+\varepsilon B\right).$$

因为 S 是有限维的,故有

$$\bigcup_{\beta>0}\bigcap_{\lambda\in(0,\beta)}\left(\frac{1}{\lambda}(S-\boldsymbol{x}')+\varepsilon B\right)\subset T_S^D(\boldsymbol{x}')+\varepsilon B.$$

据此,由定义 5.2.1 知

$$T_S^C(\boldsymbol{x})\subset\bigcap_{\varepsilon>0}\bigcup_{\alpha>0}\bigcap_{\boldsymbol{x}'\in B_S(\boldsymbol{x},\alpha)}(T_S^D(\boldsymbol{x}')+\varepsilon B)=\liminf_{\boldsymbol{x}'\to_S\boldsymbol{x}} T_S^D(\boldsymbol{x}),$$

于是结论得证. ▯

定理 5.2.4 设集合 $S_i\subset\mathscr{B}$ 非空 $(i\in I=\{1,2,\cdots\})$.

(1) 若 $S_1\subset S_2$, 则对任意的 $\boldsymbol{x}\in S_1$, 有 $T_{S_1}^D(\boldsymbol{x})\subset T_{S_2}^D(\boldsymbol{x})$.

(2) 若 $S=\bigcup_{i\in I}S_i$, 则

$$\bigcup_{i\in I}T_{S_i}^D(\boldsymbol{x})\subset T_S^D(\boldsymbol{x})\quad\forall\boldsymbol{x}\in S.$$

若 I 是有限集，则上式成为等式.

(3) 若 $S=\bigcap_{i\in I}S_i$，则

$$T_S^D(\boldsymbol{x})\subset\bigcap_{i\in I}T_{S_i}^D(\boldsymbol{x})\quad\forall\boldsymbol{x}\in S.$$

(4) 若 $I=\{1,\cdots,m\}$，$S=\prod_{i\in I}S_i$，则对任意的 $\boldsymbol{x}=(\boldsymbol{x}^1,\cdots,\boldsymbol{x}^m)\in S$，有

$$T_S^D(\boldsymbol{x})\subset\prod_{i\in I}T_{S_i}^D(\boldsymbol{x}^i),\quad T_S^C(\boldsymbol{x})=\prod_{i\in I}T_{S_i}^C(\boldsymbol{x}^i).$$

证明　由定义 5.2.1 和定义 5.2.2 容易推得结论. ▯

设 $\mathscr{D}$ 是 Banach 空间.

定理 5.2.5　设 $\Omega\subset\mathscr{B}$ 是开集. 若映射 $\varphi:\Omega\to\mathscr{D}$ 是严格可微的，集合 $S\subset\Omega$，则对任意的 $\boldsymbol{x}\in S$ 有

$$\langle D_S\varphi(\boldsymbol{x}),T_S^D(\boldsymbol{x})\rangle\subset T_{\varphi(S)}^D(\varphi(\boldsymbol{x})).$$

证明　设 $\boldsymbol{d}\in T_S^D(\boldsymbol{x})$，根据注 5.2.1 的(2)，存在序列 $\{\boldsymbol{d}^k\}\subset\mathscr{B}$，$\boldsymbol{d}^k\to\boldsymbol{d}$，$\{t_k\}\subset(0,+\infty)$，$t_k\to0^+\quad(k\to\infty)$，使得对任何 $k=1,2,\cdots$，有 $\boldsymbol{x}+\lambda_k\boldsymbol{d}^k\in S$. 由 φ 是严格可微的，根据定义 4.2.3 有

$$\boldsymbol{x}^k=\frac{\varphi(\boldsymbol{x}+t_k\boldsymbol{d}^k)-\varphi(\boldsymbol{x})}{t_k}\to\langle D_S\varphi(\boldsymbol{x}),\boldsymbol{d}\rangle,$$

从而 $\varphi(\boldsymbol{x})+\lambda_k\boldsymbol{x}^k=\varphi(\boldsymbol{x}+t_k\boldsymbol{d}^k)\in\varphi(S)$. 由此，得到 $\langle D_S\varphi(\boldsymbol{x}),\boldsymbol{d}\rangle\in T_{\varphi(S)}^D(\varphi(\boldsymbol{x}))$. ▯

注 5.2.7　特别地，当 $\varphi:\mathscr{B}\to\mathscr{D}$ 是线性映射时，我们有 $\varphi(T_S^D(\boldsymbol{x}))\subset T_{\varphi(S)}^D(\varphi(\boldsymbol{x}))$.

定理 5.2.6　设集合 $S\subset\mathscr{B}$，$M\subset\mathscr{D}$. 若 φ 是 S 的一个开邻域到 $\mathscr{D}$ 的严格可微映射，

$$V=\{\boldsymbol{x}\in S\,|\,\varphi(\boldsymbol{x})\in\mathscr{D}\}=S\cap\varphi^{-1}(M),$$

则

$$T_V^D(\boldsymbol{x}) \subset T_S^D(\boldsymbol{x}) \cap D_S\varphi(\boldsymbol{x})^{-1}T_M^D\big(\varphi(\boldsymbol{x})\big).$$

证明 由定理 5.2.5 有 $T_V^D(\boldsymbol{x}) \subset T_S^D(\boldsymbol{x})$. 再据定理 5.2.6 得

$$\langle D_S\varphi(\boldsymbol{x}), T_V^D(\boldsymbol{x}) \rangle \subset T_{\varphi(S)}^D\big(\varphi(\boldsymbol{x})\big) \subset T_M^D\big(\varphi(\boldsymbol{x})\big).$$

因此有 $T_V^D(\boldsymbol{x}) \subset D_S\varphi(\boldsymbol{x})^{-1}T_M^D\big(\varphi(\boldsymbol{x})\big)$. □

定理 5.2.7 设集合 $S \subset \Omega \subset R^n$, $M \subset R^m$, 并且 Ω 是开集, S 和 M 是闭集. 又设映射 $\varphi:\Omega \to R^m$ 是严格可微的, $V = \{\boldsymbol{x} \in S \mid \varphi(\boldsymbol{x}) \in M\} = S \cap \varphi^{-1}(M)$. 若存在 $\boldsymbol{x} \in V$, 使得

$$\langle D_S\varphi(\boldsymbol{x}), T_S^C(\boldsymbol{x}) \rangle - T_M^D\big(\varphi(\boldsymbol{x})\big) = R^n,$$

则

$$T_S^C(\boldsymbol{x}) \cap D_S\varphi(\boldsymbol{x})^{-1}T_M^C\big(\varphi(\boldsymbol{x})\big) \subset T_V^C(\boldsymbol{x}).$$

证明 作集值映射 $\psi:R^n \to 2^{R^m}$, $x \mapsto \psi(\boldsymbol{x})$,

$$\psi(\boldsymbol{x}) = \begin{cases} \varphi(\boldsymbol{x}) - M, & \boldsymbol{x} \in S; \\ \varnothing, & \boldsymbol{x} \notin S. \end{cases}$$

显然, 有 $\psi^{-1}(\boldsymbol{0}) = V$. 已知 $\boldsymbol{x} \in V$, 则存在 $\boldsymbol{x}$ 的一个开邻域 $U(\boldsymbol{x})$ 以及 $\alpha > 0$ 和 $\beta > 0$, 使得

$$d_{\psi^{-1}(\boldsymbol{0})}(\boldsymbol{x}') \leqslant \beta \|\boldsymbol{y}\| \quad \forall \boldsymbol{y} \in \alpha B, \quad \forall \boldsymbol{x}' \in \psi^{-1}(\boldsymbol{y}) \cap U(\boldsymbol{x}).$$

选择充分小的 α, 使得

$$\|\varphi(\boldsymbol{x}') - \varphi(\boldsymbol{x})\| < \alpha,$$

$$S \cap U(\boldsymbol{x}) \subset \psi^{-1}\Big(\varphi(\boldsymbol{x}') - \pi_M\big(\varphi(\boldsymbol{x}')\big)\Big),$$

其中 $\pi_M\big(\varphi(\boldsymbol{x}')\big) = \big\{\varphi(\boldsymbol{x}') \in M \,\big|\, \|\varphi(\boldsymbol{x}') - \boldsymbol{y}\| = d_M(\boldsymbol{y}),\ \boldsymbol{x}' \in R^n\big\}$,

则 $$\Big\|\varphi(\boldsymbol{x}') - \pi_M\big(\varphi(\boldsymbol{x}')\big)\Big\| \leqslant \big\|\varphi(\boldsymbol{x}') - \varphi(\boldsymbol{x})\big\| < \alpha.$$

由此, 对任意的 $\boldsymbol{x}' \in S \cap U(\boldsymbol{x})$ 有

$$d_{\psi^{-1}(\boldsymbol{0})}(\boldsymbol{x}') \leqslant \beta\Big\|\varphi(\boldsymbol{x}') - \pi_M\big(\varphi(\boldsymbol{x}')\big)\Big\| = \beta d_M\big(\varphi(\boldsymbol{x}')\big). \tag{5.2.6}$$

下面设 $d \in T_S^C(x) \cap D_S\varphi(x)^{-1}T_M^C\big(\varphi(x)\big)$，则存在 $\varepsilon > 0$，$\delta > 0$ 使得当 $\|x' - x\| \leqslant \varepsilon$ 和 $\lambda \leqslant \delta$ 时，有 $x + \lambda d \in S \cap U(x)$. 从 (5.2.6)我们有

$$\frac{1}{\lambda} d_{\varphi^{-1}(0)}(x' + \lambda d) \leqslant \beta \frac{d_M\big(\varphi(x' + \lambda d)\big)}{\lambda}$$

$$\leqslant \beta \frac{d_S\big(\varphi(x') + \lambda \langle D_S\varphi(x), d\rangle\big)}{\lambda}$$

$$+ \beta \frac{\big\|\varphi(x' + \lambda d) - \varphi(x') - \lambda\langle D_S\varphi(x), d\rangle\big\|}{\lambda}.$$

因为 $\langle D_S\varphi(x), d\rangle \in T_M^C\big(\varphi(x)\big)$，所以上式右端两式当 $\lambda \to 0$ 时趋于 0，从而得 $d \in T_V^C(x)$. ▯

由定理 5.2.7，容易推知以下三个推论成立.

推论 5.2.8　设 $\varphi: R^n \to R^m$ 是连续严格可微映射，$M \subset \mathscr{D}$ 是闭集，$\varphi(x) \in M$. 若 $\mathrm{Im} D_S\varphi(x) - T_M^C\big(\varphi(x)\big) = M$，则

$$D_S\varphi(x)^{-1}T_M^C\big(\varphi(x)\big) \subset T_{\varphi^{-1}(M)}^C(x).$$

推论 5.2.9　设 $S_1, S_2 \subset R^n$ 是非空闭集，$x \in S_1 \cap S_2$. 若 $T_{S_1}^C(x) - T_{S_2}^C(x) = R^n$，则

$$T_{S_1}^C(x) \cap T_{S_2}^C(x) \subset T_{S_1 \cup S_2}(x).$$

推论 5.2.10　设 $S_i \subset R^n$ 是非空闭集 $(i = 1, \cdots, m)$，$x \in \bigcap\limits_{i=1}^m S_i$. 若对任意的 $d^1, \cdots, d^m \in R^n$ 有 $\bigcap\limits_{i=1}^m \big(T_{S_i}^C(x) - d^i\big) \neq \varnothing$，则

$$\bigcap_{i=1}^m T_{S_i}^C(x) \subset T^C_{\bigcap\limits_{i=1}^m S_i}(x).$$

§ 5.3　D-切导数

从这一节开始，我们将利用切锥来定义导数和次微分. 设 $\mathscr{B}$

是 Banach 空间，$f:\mathscr{B}\to R\cup\{+\infty\}$ 是正常函数. 作集值映射 $f_+:\mathscr{B}\to 2^R$，$\boldsymbol{x}\mapsto R$，

$$f_+(\boldsymbol{x})=\begin{cases}f(\boldsymbol{x})+R_+, & f(\boldsymbol{x})<+\infty;\\ \varnothing, & f(\boldsymbol{x})=+\infty.\end{cases}$$

定义 5.3.1 设 $f:\mathscr{B}\to R\cup\{+\infty\}$ 是正常函数，点 $\boldsymbol{x}^0\in\mathscr{B}$，$\boldsymbol{d}\in\mathscr{B}$，$T^D_{\mathrm{epi}f_+}\big(\boldsymbol{x}^0, f(\boldsymbol{x}^0)\big)$ 是 $\mathrm{epi}f_+$ 在点 $\big(\boldsymbol{x}^0, f(\boldsymbol{x}^0)\big)$ 处的 D-切锥. 又设 $(\boldsymbol{d}, \eta)\in T^D_{\mathrm{epi}f_+}\big(\boldsymbol{x}^0, f(\boldsymbol{x}^0)\big)$，记

$$f^D(\boldsymbol{x}^0;\boldsymbol{d})=\inf\Big\{\eta\,\Big|\,(\boldsymbol{d},\eta)\in T^D_{\mathrm{epi}f_+}\big(\boldsymbol{x}^0, f(\boldsymbol{x}^0)\big)\Big\},\tag{5.3.1}$$

则称 $f^D(\boldsymbol{x}^0;\boldsymbol{d})$是函数 f 在点 $\boldsymbol{x}^0$ 处沿方向 $\boldsymbol{d}$ 的 D-切导数.

注 5.3.1 在 $T^D_{\mathrm{epi}f_+}\big(\boldsymbol{x}, f(\boldsymbol{x})\big)$ 中对每一给定的$(\boldsymbol{d},\eta)$构成的集合或为 R，或为$(\eta, +\infty)$，或为空集.

由于 D-切锥不一定是凸的，故不用它来定义相应的次梯度. 因为即使给出次梯度的定义，也得不到多少好的性质. 下面研究 D-切导数的一些性质.

定理 5.3.1 设 $f:\mathscr{B}\to R\cup\{+\infty\}$ 是正常函数，则

$$f^D(\boldsymbol{x};\boldsymbol{d})=\liminf_{\substack{\boldsymbol{d}'\to\boldsymbol{d}\\ t\to 0^+}}\frac{f(\boldsymbol{x}+t\boldsymbol{d}')-f(\boldsymbol{x})}{t}.$$

并且，若对任意的 $\boldsymbol{d}\in\mathscr{B}$ 有 $f^D(\boldsymbol{x};\boldsymbol{d})<+\infty$，则 $f^D(\boldsymbol{x};\boldsymbol{d})$关于 $\boldsymbol{d}$ 是正齐次和下半连续的.

证明 设 $(\boldsymbol{d}, \eta)\in T^D_{\mathrm{epi}f_+}\big(\boldsymbol{x}, f(\boldsymbol{x})\big)$，则对任意的 $\varepsilon_1>0$ 和 $\varepsilon_2>0$，以及任意的 $\alpha>0$，存在 $\boldsymbol{d}'\in\boldsymbol{d}+\varepsilon_2 B$ 和 $t\in[0, \alpha]$，使得

$$\eta\in\frac{f_+(\boldsymbol{x}+t\boldsymbol{d}')-f(\boldsymbol{x})}{t}+\varepsilon_1 B',$$

其中 $B'=[-1, 1]$. 从上式可推得

$$\varepsilon_1+\eta\geqslant\frac{f(\boldsymbol{x}+t\boldsymbol{d}')-f(\boldsymbol{x})}{t}$$

$$\geqslant \inf_{t<\alpha} \inf_{\|\boldsymbol{d}'-\boldsymbol{d}\|\leqslant\varepsilon_2} \frac{f(\boldsymbol{x}+t\boldsymbol{d}')-f(\boldsymbol{x})}{t},$$

因此

$$\eta \geqslant \liminf_{\substack{\boldsymbol{d}'\to\boldsymbol{d}\\ t\to 0^+}} \frac{f(\boldsymbol{x}+t\boldsymbol{d}')-f(\boldsymbol{x})}{t} - \varepsilon_1.$$

设 $a = \liminf_{\substack{\boldsymbol{d}'\to\boldsymbol{d}\\ t\to 0^+}} \frac{f(\boldsymbol{x}+t\boldsymbol{d}')-f(\boldsymbol{x})}{t} - \varepsilon_1$，则有 $a \leqslant f^D(\boldsymbol{x};\boldsymbol{d})$.

另一方面，对任意的 $M > a$，有

$$\sup_{\substack{\alpha>0\\ \delta>0}} \inf_{t<\alpha} \inf_{\|\boldsymbol{d}'-\boldsymbol{d}\|\leqslant\delta} \frac{f(\boldsymbol{x}+t\boldsymbol{d}')-f(\boldsymbol{x})}{t} < M,$$

即对任意的 $\alpha > 0$，$\delta > 0$，存在 $t < \alpha$ 和 $\boldsymbol{d}' \in \boldsymbol{d} + \delta B$，使得

$$\frac{f(\boldsymbol{x}+t\boldsymbol{d}')-f(\boldsymbol{x})}{t} \leqslant M.$$

由此，得 $M \in \frac{f_+(\boldsymbol{x}+t\boldsymbol{d}')-f(\boldsymbol{x})}{t}$，也就是说 $(\boldsymbol{d}, a) \in T^D_{\mathrm{epi} f_+}(\boldsymbol{x}, f(\boldsymbol{x}))$. 再由定义 5.3.1 便知 $a = f^D(\boldsymbol{x};\boldsymbol{d})$. 其余结论容易推得. □

注 5.3.2　由定理 5.3.1 可知，若 $f: R^n \to R$ 在 $\boldsymbol{x} \subset R^n$ 处可微，则

$$f^D(\boldsymbol{x};\boldsymbol{d}) = \langle \nabla f(\boldsymbol{x}), \boldsymbol{d} \rangle.$$

若 f 是凸函数，则对任意的 $\boldsymbol{d} \in R^n$ 有

$$f^D(\boldsymbol{x};\boldsymbol{d}) = \liminf_{\boldsymbol{d}'\to\boldsymbol{d}} \left(\inf_{t>0} \frac{f(\boldsymbol{x}+t\boldsymbol{d}')-f(\boldsymbol{x})}{t} \right).$$

定理 5.3.2　设 $f: \mathscr{B} \to R \cup \{+\infty\}$ 是正常函数. 若 f 在 $\boldsymbol{x} \in \mathrm{int}(\mathrm{dom} f)$ 处是 Lipschitz 的，则

$$f^D(\boldsymbol{x};\boldsymbol{d}) = \liminf_{t\to 0^+} \frac{f(\boldsymbol{x}+t\boldsymbol{d})-f(\boldsymbol{x})}{t} \quad \forall \boldsymbol{d} \in \mathscr{B},$$

并且 $f^D(\boldsymbol{x};\boldsymbol{d})$有限.

证明 由定理 5.3.1 的证明，可知结论成立. ▯

定理 5.3.3 设 $f:\mathscr{B}\to R\cup\{+\infty\}$ 是正常函数，集合 $S\subset\mathscr{B}$ 非空. 记

$$f|_S(\boldsymbol{x})=\begin{cases}f(\boldsymbol{x}), & \boldsymbol{x}\in S;\\ +\infty & \boldsymbol{x}\notin S,\end{cases}$$

则

$$f^D(\boldsymbol{x};\boldsymbol{d})\leqslant (f|_S)^D(\boldsymbol{x};\boldsymbol{d})\quad \forall \boldsymbol{x}\in\mathscr{B},\ \forall \boldsymbol{d}\in T_S^D(\boldsymbol{x}).$$

若 f 在 $\boldsymbol{x}\in S$ 处是严格可微的，则有

$$(f|_S)^D(\boldsymbol{x};\boldsymbol{d})=\begin{cases}\langle D_S f(\boldsymbol{x}),\boldsymbol{d}\rangle, & \boldsymbol{d}\in T_S^D(\boldsymbol{x});\\ +\infty, & \boldsymbol{d}\notin T_S^D(\boldsymbol{x}).\end{cases}$$

证明 由定义 5.3.1，易知结论成立. ▯

定理 5.3.4 设 $f:\mathscr{B}\to R\cup\{+\infty\}$ 是正常函数. 若 $\boldsymbol{x}^0\in\mathrm{dom}f$，并且对任意的 $\boldsymbol{x}\in\mathscr{B}$ 有 $f(\boldsymbol{x}^0)\leqslant f(\boldsymbol{x})$，则

$$f^D(\boldsymbol{x};\boldsymbol{d})\geqslant 0\quad \forall \boldsymbol{d}\in\mathscr{B}.$$

证明 由定理 5.3.1 直接可得. ▯

以下为 ε-变分原理.

定理 5.3.5 设 $f:\mathscr{B}\to R\cup\{+\infty\}$ 是下半连续和有界的正常函数，$\boldsymbol{x}^0\in\mathrm{dom}f$，则对任意的 $\varepsilon>0$ 存在 $\boldsymbol{x}_\varepsilon\in\mathrm{dom}f$，使得

$$f(\boldsymbol{x}_\varepsilon)+\varepsilon\|\boldsymbol{x}_\varepsilon-\boldsymbol{x}^0\|\leqslant f(\boldsymbol{x}^0),\tag{5.3.2}$$

$$0\leqslant f^D(\boldsymbol{x}_\varepsilon;\boldsymbol{d})+\varepsilon\|\boldsymbol{d}\|\quad \forall \boldsymbol{d}\in\mathscr{B}.\tag{5.3.3}$$

证明 根据文献[16]的定理 5.3.1，得知存在 $\boldsymbol{x}_\varepsilon\in\mathrm{dom}f$ 满足(5.3.2)，并且

$$f(\boldsymbol{x}_\varepsilon)=\min\{f(\boldsymbol{x})+\varepsilon\|\boldsymbol{x}_\varepsilon-\boldsymbol{x}\|\ |\boldsymbol{x}\in\mathscr{B}\}.$$

设 $\boldsymbol{d}\in\mathrm{dom}(f^D(\boldsymbol{x}_\varepsilon;\cdot))$，则对任意的 $\eta>0$，$\delta>0$，$\alpha>0$，存在 $t\leqslant\alpha$ 和 $\boldsymbol{d}'\in\boldsymbol{d}+\delta B$，使得

$$\frac{f(x_\varepsilon + td') - f(x_\varepsilon)}{t} \leqslant f^D(x_\varepsilon; d) + \eta.$$

再由文献[16]的定理 5.3.1 可得

$$-\varepsilon\delta - \varepsilon \| d \| \leqslant - \varepsilon \| d \| \leqslant \frac{f(x_\varepsilon + td') - f(x_\varepsilon)}{t},$$

因此有

$$0 \leqslant f^D(x_\varepsilon; d) + \varepsilon \| x \| + \varepsilon\delta + \eta.$$

令 δ, $\eta \to 0$，即得(5.3.3)成立. ▯

定理 5.3.6 设 $f_1, f_2: \mathscr{B} \to R \cup \{+\infty\}$ 是正常函数，则

$$f_1^D(x; d) + f_2^D(x; d) \leqslant (f_1 + f_2)^D(x; d).$$

证明 由定义 5.3.1 直接可得到. ▯

定理 5.3.7 设 $\mathscr{D}$ 是 Banach 空间，$f: \mathscr{D} \to R \cup \{+\infty\}$ 是正常函数，$\varphi: \mathscr{B} \to \mathscr{D}$ 是严格可微映射. 若 $\varphi(x) \in \mathrm{dom} f$，则

$$f^D\left(\varphi(x); \langle D_S\varphi(x), d \rangle\right) \leqslant \left(f(\varphi)\right)^D(x; d) \quad \forall d \in \mathscr{B}.$$

证明 由定理 5.2.6 可得到. ▯

§ 5.4 C-切导数和 C-切微分

本节利用函数上图象的 C-切锥来定义导数. 由于 C-切锥是凸的，因而还可以进而引入相应的次梯度和次微分. 我们将看到，当函数满足 Lipschitz 条件时，本节定义的导数与第 4 章引进的 G-方向导数是一致的.

设 $\mathscr{B}$ 是 Banach 空间.

定义 5.4.1 设 $f: \mathscr{B} \to R \cup \{+\infty\}$ 是正常函数，点 $x^0 \in \mathscr{B}$，$d \in \mathscr{B}$，$T^C_{\mathrm{epi} f_+}(x^0, f(x^0))$ 是 $\mathrm{epi} f_+$ 在点 $(x^0, f(x^0))$ 处的 C-切锥. 又设 $(d, \eta) \in T^C_{\mathrm{epi} f_+}\left(x^0, f(x^0)\right)$，记

$$f^C(x^0; d) = \inf\{\eta \mid (d, \eta) \in T_{\mathrm{epi}f_+}(x^0, f(x^0))\}, \tag{5.4.1}$$

则称 $f^C(x^0; d)$是 f 在点 x^0 处沿 d 的上图象导数，或 C-切导数.

定理 5.4.1 设 $x \in \mathscr{B}$. 若对任意的 $d \in \mathscr{B}$, $f^C(x; d) > -\infty$，则 $f^C(x; d)$关于 d 是正齐次和下半连续的凸函数.

证明 由 C-切锥的定义和性质即可得到. ▯

设 $\mathscr{B}^*$是 $\mathscr{B}$ 的对偶空间.

定义 5.4.2 设 $f:\mathscr{B} \to R \cup \{+\infty\}$ 是正常函数，点 $x^0 \in \mathscr{B}$，则称集合

$$\partial^C f(x^0) = \{x^* \in \mathscr{B}^* \mid \langle x^*, d\rangle \leqslant f^C(x^0; d) \quad \forall d \in \mathscr{B}\} \tag{5.4.2}$$

是 f 在点 x^0 处的 C-切微分.

注 5.4.1 若存在 $d \in \mathscr{B}$ 使得 $f^C(x; d) = -\infty$，则 $\partial^C f(x) = \varnothing$. 若 f 在点 x 处严格可微，则 $f^C(x; d) = \langle D_S f(x), d\rangle$，$\partial^C f(x) = \{D_S f(x)\}$. 当 f 是凸函数时，这里定义的 C-次微分与第 3 章定义的次微分是一致的.

定理 5.4.2 设 $f:\mathscr{B} \to R$ 是实值函数，$x \in \mathrm{int}(\mathrm{dom}f)$. 若 f 在 x 处是 Lipschitz 的，则

$$f^C(x; d) = f^0(x; d).$$

证明 由定理 4.4.14 可以推得. ▯

设 $x, x' \in \mathscr{B}$, $\alpha \in R$，记号$(x', \alpha) \downarrow x$ 表示 $\alpha \geqslant f(x')$, $x' \to x$, $\alpha \to f(x)$.

定理 5.4.3 设 $f:\mathscr{B} \to R \cup \{+\infty\}$ 是正常函数. 若 $x \in \mathrm{dom}f$, $\alpha \in R$，则

$$f^C(x; d) = \lim_{\substack{(x', \alpha)\downarrow x \\ t \to 0^+}} \sup \inf_{d' \to d} \frac{f(x' + td') - \alpha}{t}$$

即

$$f^C(x;d)=\lim_{\substack{\varepsilon\to0^+,(x',\alpha)\downarrow x\\ t\to0^+}}\sup\inf_{d'\in d+\varepsilon B}\frac{f(x'+td')-\alpha}{t}.$$

证明　取 $(d,\eta)\in T^C_{\mathrm{epi}f_+}(x,f(x))$，任给 $\varepsilon>0$，设 $(x^k,\alpha_k)\downarrow x$，$t_k\to0^+$ $(k\to\infty)$，则存在 $(d^k,\eta_k)\to(d,\eta)$ 使得 $(x^k,\alpha_k)+t_k(d^k,\eta_k)\in\mathrm{epi}f_+$，即

$$\alpha_k+t_k\eta_k\geqslant f(x^k+t_kd^k).$$

由此，我们有

$$\inf_{d'\to d}\frac{f(x^k+t_kd')-\alpha_k}{t}\leqslant\frac{f(x^k+t_kd^k)-\alpha_k}{t}\leqslant\eta_k,$$

从而得 $f^C(x;d)\leqslant\eta$.

反之，设 $f^C(x;d)\leqslant\eta$，$(x^k,\alpha_k)\downarrow x$，$t_k\to0^+$ $(k\to\infty)$，则对每一个 $k=1,2,\cdots$，存在一个 $\varepsilon_k<1/k$，使得

$$\lim_{\substack{(x^k,\alpha)\downarrow x\\ t_k\to0^+}}\sup\inf_{d'\to d}\frac{f(x^k+t_kd')-\alpha}{t_k}\leqslant\eta+\frac{1}{k}.$$

从上式有

$$\inf_{d'\in d+\varepsilon_kB}\frac{f(x^k+t_kd')-\alpha_k}{t_k}\leqslant\eta+\frac{2}{k},$$

故对每一 ε_k，存在 $d^k\in d+\varepsilon_kB$ 使得

$$\frac{f(x^k+t_kd^k)-\alpha_k}{t_k}\leqslant\eta+\frac{3}{k}.$$

由此，定义 $\eta_k=\max\left\{\eta,\dfrac{f(x^k+t_kd^k)-\alpha_k}{t_k}\right\}(k=1,2,\cdots)$，有 $\eta_k\to\eta(k\to\infty)$，故

$$\alpha_k+t_k\eta_k\geqslant f(x^k+t_kd^k),\ k=1,2,\cdots.$$

由定义 5.2.2 知 $(d,\eta)\in T^C_{\mathrm{epi}f_+}(x,f(x))$，再据定义 5.4.1 便知

结论成立.□

推论 5.4.4 设 $f:\mathscr{B}\to R\cup\{+\infty\}$ 是正常函数. 若 $x\in \mathrm{dom}f$, 则 $\partial^C f(x)=\varnothing$ 当且仅当 $f^C(x;d)=-\infty$. 否则, 有

$$\partial^C f(x)=\{x^*\in\mathscr{B}^*\mid\langle x^*,d\rangle\leqslant f^C(x;d)\quad\forall d\in\mathscr{B}\}$$

和

$$f^C(x;d)=\sup\{\langle x^*,d\rangle\mid x^*\in\partial^C f(x)\}.$$

证明 由定义 5.4.1、定义 5.4.2 和定理 5.4.3 可得到.□

§5.5 H-切导数

这一节,通过§4.4 中的 H-切锥来引进另一种切导数. 它与 C-切导数有着紧密的联系. 在§5.6 中将再进一步讨论 C-切导数的有关问题.

定义 5.5.1 设 $f:\mathscr{B}\to R\cup\{+\infty\}$ 是实值函数, 点 $x^0\in\mathscr{B}$, $d\in\mathscr{B}$, $T^H_{\mathrm{epi}f}(x^0,f(x^0))$ 是 $\mathrm{epi}f$ 在点 $(x^0,f(x^0))$ 处的 H-切锥. 称

$$f^H(x;d)=\inf\{\eta\mid(d,\eta)\in T^H_{\mathrm{epi}f}(x^0,f(x^0))\}\qquad(5.5.1)$$

是函数 f 在点 x^0 处沿方向 d 的 H-切导数. 若 $f^H(x^0;d)<+\infty$, 则称 f 在点 x^0 处关于 d 是有限 H-可切导的. 设 $x\in\mathscr{B}$, 若函数 $f^H:\mathscr{B}\times\mathscr{B}\to R$, $(x,d)\mapsto f^H(x;d)$ 关于 d 是正常的(即存在 $\bar{d}$ 使得 $f^H(x;\bar{d})<+\infty$), 则称 f 在点 x 处关于 d 是正常 H-可切导的.

注 5.5.1 显然, 若 f 在 $x\in\mathscr{B}$ 处是 Lipschitz 的, 则 f 在 x 处关于任意的 $d\in\mathscr{B}$ 是正常 H-可切导的. 若 $d\in\mathrm{dom}f^H(x;\cdot)$, 则 f 在 x 处关于 d 是有限 H-可切导的.

注 5.5.2 设 $x\in\mathscr{B}$, 由定义 5.3.1、定义 5.4.1 和定义 5.5.1, 可推知有

$$f^D(\boldsymbol{x};\boldsymbol{d}) \leqslant f^C(\boldsymbol{x};\boldsymbol{d}) \leqslant f^H(\boldsymbol{x};\boldsymbol{d}) \quad \forall \boldsymbol{d} \in \mathscr{B}.$$

定理 5.5.1　设 $f:\mathscr{B}\to R\cup\{+\infty\}$ 是实值函数，$\boldsymbol{x}\in \mathrm{dom} f$. 若 $f^H(\boldsymbol{x};\boldsymbol{d})<+\infty$，则

$$f^H(\boldsymbol{x};\boldsymbol{d}) = \lim_{\substack{(\boldsymbol{x}',\alpha)\downarrow \boldsymbol{x} \\ t\to 0^+}} \sup_{\boldsymbol{d}'\to \boldsymbol{d}} \frac{f(\boldsymbol{x}'+t\boldsymbol{d}')-\alpha}{t}.$$

证明　设 $(\boldsymbol{d},\eta)\in T^H_{\mathrm{epi}f_+}(\boldsymbol{x},f(\boldsymbol{x}))$，根据定义 4.4.3，存在 $\varepsilon>0$ 有

$$\big((\boldsymbol{x},f(\boldsymbol{x}))+\varepsilon B_{\mathscr{B}\times R}\big)\cap\big(\mathrm{epi}f_+ + t\big((\boldsymbol{d},\eta)+\varepsilon B_{\mathscr{B}\times R}\big)\big)$$
$$\subset \mathrm{epi}f_+,\ t\in(0,\varepsilon). \tag{5.5.2}$$

设 $(\boldsymbol{x}',\alpha)\downarrow \boldsymbol{x}$，$\boldsymbol{d}'\to\boldsymbol{d}$，从上式得 $\alpha+t\eta\geqslant f(\boldsymbol{x}'+t\boldsymbol{d}')$，即

$$\frac{f(\boldsymbol{x}'+t\boldsymbol{d}')-\alpha}{t}\leqslant\eta.$$

因此，由(5.5.1)有 $f^H(\boldsymbol{x};\boldsymbol{d})\leqslant\eta$.

反之，设 $f^H(\boldsymbol{x};\boldsymbol{d})\leqslant\eta$，由定义 5.5.1 可知，对所有的 $(\boldsymbol{x}',\alpha)\downarrow\boldsymbol{x}$，$\boldsymbol{d}'\to\boldsymbol{d}$，$t\to0^+$，有

$$\frac{f(\boldsymbol{x}'+t\boldsymbol{d}')-\alpha}{t}\leqslant\eta.$$

这说明存在 $\varepsilon>0$ 使得(5.5.2)成立，于是 $(\boldsymbol{d},\eta)\in T^H_{\mathrm{epi}f_+}(\boldsymbol{x},f(\boldsymbol{x}))$. □

定理 5.5.2　设 $f:\mathscr{B}\to R\cup\{+\infty\}$ 是正常函数. 若下面 5 个条件之一成立，则对任意的 $\boldsymbol{d}\in\mathscr{B}$，$f$ 在 $\boldsymbol{x}$ 处关于 $\boldsymbol{d}$ 是有限 H-可切导的：

(1) f 在 $\boldsymbol{x}$ 处是 Lipschitz 的.

(2) $f=\delta_{\mathrm{epi}f_+}(\boldsymbol{x})$，并且 $\mathrm{epi}f_+$ 在 $\boldsymbol{x}$ 处存在超切向量.

(3) f 是凸函数，并且在某点(不必是 $\boldsymbol{x}$)的邻域有界.

(4) f 关于由某个内部非空的闭凸锥 $K\in\mathscr{B}$ 确定的偏序是非减函数(即由 $\boldsymbol{x}_2-\boldsymbol{x}_1\in K$，有 $f(\boldsymbol{x}_1)\leqslant f(\boldsymbol{x}_2)$).

(5) $\mathscr{B}=R^n$. f 在 x 的一个邻域内是上半连续的，并且 $\partial^C f(x)\neq\varnothing$.

证明 (1)(2) 由定理 5.5.1 即可得到.

(3) 记 $C=\mathrm{epi}f_+$. 设 f 在 $x+td$ 的某个邻域 U 内有界，则对某个 $t_0>0$，存在 $\beta\in(1,+\infty)$ 有

$$|f(x')|<\beta-1 \quad \forall x'\in U.$$

因此，有 (d,α) 使得 $\big(x,f(x)\big)+t_0(d,x)\in\mathrm{int}C$，其中 $\alpha=\dfrac{\beta-f(x)}{t_0}$. 记 $B=B_{\mathscr{B}}\times B_R$，可以选择 $\delta>0$，使得

$$\Big(\big(x,f(x)\big)+\delta B\Big)+t_0\big((d,\alpha)+\delta B\big)\subset C.$$

已知 f 是凸的，容易验证 C 也是凸的. 由上式知对任意的 $t\in(0,t_0)$，$x'\in x+\delta B_{\mathscr{B}}$，$d'\in d+\delta B_{\mathscr{B}}$ 和 $f(x')\in f(x)+\delta B_R$，有 r'，$r''\in\alpha+\delta B_R$ 使得 $f(x')\leqslant r'$ 和 $f(x'+t_0d')\leqslant r''$ 成立，并且有

$$x'+tx=\left(1-\frac{t}{t_0}\right)x'+\frac{t}{t_0}(x'+t_0d'),$$

$$\begin{aligned}f(x'+td')&\leqslant\left(1-\frac{t}{t_0}\right)f(x')+\frac{t}{t_0}f(x'+t_0d')\\&\leqslant\left(1-\frac{t}{t_0}\right)r'+\frac{t}{t_0}(r'+t_0r'')\leqslant r'+tr''.\end{aligned}$$

据此，得到

$$\Big(\big(x,f(x)\big)+\delta B\Big)\cap C+t\big[(d,\alpha)+\delta B\big]\subset C$$

$$\forall t\in(0,t_0),$$

故有 $(d,\alpha)\in T^H_{\mathrm{epi}f_+}\big(x,f(x)\big)$. 由定理 5.5.1，即知 $f^H(x;d)$ 存在.

(4) 从已知可设 $-d\in\mathrm{int}K$ 使得 $-d+\varepsilon B\subset K$ (其中 $\varepsilon>0$). 于是，对 $-p\in-d+\varepsilon B$ 和 $t>0$ 有 $-tp\in K$. 因为 f 关于

K 是非减的，故对任意的 $\boldsymbol{y}$，由 $\boldsymbol{y}-(\boldsymbol{y}+t\boldsymbol{p})\in K$ 有 $f(\boldsymbol{y}+t\boldsymbol{p})\leqslant f(\boldsymbol{y})$，从而得

$$\frac{f(\boldsymbol{y}+t\boldsymbol{p})-f(\boldsymbol{y})}{t}\leqslant 0.$$

因此，当 $\boldsymbol{y}\to\boldsymbol{x}$，$\boldsymbol{p}\to\boldsymbol{d}$ 和 $t\to 0^+$ 时，上式有界，也即 $f^H(\boldsymbol{x};\boldsymbol{d})<+\infty$.

(5) 由 $\partial^C f(\boldsymbol{x})\neq\varnothing$，令 $D=\{\boldsymbol{d}\mid f^C(\boldsymbol{x};\boldsymbol{d})<+\infty\}$，从推论 5.4.4 知 $\mathrm{int}D\neq\varnothing$. 因为 $\mathscr{B}$ 是有限维的，并且 $f^C(\boldsymbol{x};\boldsymbol{d})$在 $\mathrm{int}D$ 内是有界的，利用推论 4.5.11，即得 $(\boldsymbol{d},\eta)\in\mathrm{int}(\mathrm{epi}f^C(\boldsymbol{x};\boldsymbol{d}))=\mathrm{int}T^C_{\mathrm{epi}f_+}(\boldsymbol{x},f(\boldsymbol{x}))=T^H_{\mathrm{epi}f_+}(\boldsymbol{x},f(\boldsymbol{x}))$. □

§5.6 若干关系

现在继续讨论 C-切导数和 C-切微分的性质，也讨论与 D-切导数有关的性质.

设 $\mathscr{B}$ 是 Banach 空间，$\mathscr{B}^*$ 是 $\mathscr{B}$ 的对偶空间，记

$$\mathrm{dom}f^C(\boldsymbol{x};\ \cdot\ )=\{\boldsymbol{d}\in\mathscr{B}\mid f^C(\boldsymbol{x};\ \cdot\ )<+\infty\},$$

$$\mathrm{dom}f^H(\boldsymbol{x};\ \cdot\ )=\{\boldsymbol{d}\in\mathscr{B}\mid f^H(\boldsymbol{x};\ \cdot\ )<+\infty\}.$$

定理 5.6.1 设 $f:\mathscr{B}\to R\cup\{+\infty\}$ 是正常函数，$\boldsymbol{x}\in\mathrm{dom}f$，$\boldsymbol{d}\in\mathscr{B}$，$f$ 在 $\boldsymbol{x}$ 处关于 $\boldsymbol{d}$ 是正常 H-可切导的.

(1) $\mathrm{int\,dom}f^C(\boldsymbol{x};\ \cdot\ )=\mathrm{dom}f^H(\boldsymbol{x};\ \cdot\ )$.

(2) $f^C(\boldsymbol{x};\boldsymbol{d})=\liminf\limits_{\boldsymbol{d}'\to\boldsymbol{d}}f^H(\boldsymbol{x};\boldsymbol{d}')\,(\forall\boldsymbol{d}\in\mathrm{dom}f^C(\boldsymbol{x};\ \cdot\ ))$.

(3) f^C 在 $(\boldsymbol{x},\boldsymbol{d}')\,(\boldsymbol{d}'\in\mathrm{int\,dom}f^C(\boldsymbol{x};\ \cdot\ ))$ 处是上半连续的.

(4) f^C 关于 $\boldsymbol{d}$ 在 $\boldsymbol{d}'\in\mathrm{int\,dom}f^C(\boldsymbol{x},\ \cdot\ )$ 处是连续的.

(5) 若 $\mathrm{dom}f^H(\boldsymbol{x},\ \cdot\ )=\mathscr{B}$，则 $\partial^C(-f)(\boldsymbol{x})=-\partial^C f(\boldsymbol{x})$.

证明 (1) 设 $\boldsymbol{d}=\mathrm{dom}f^H(\boldsymbol{x};\ \cdot\ )$，由注 5.5.2 有 $f^C(\boldsymbol{x};\boldsymbol{d})\leqslant$

$f^H(\boldsymbol{x};\ \boldsymbol{d})$，所以 $\boldsymbol{d} \in \operatorname{int}\operatorname{dom}(f^C(\boldsymbol{x};\ \cdot))$. 反之，设 $\boldsymbol{d} \in \operatorname{int}\operatorname{dom} f^C(\boldsymbol{x};\ \cdot)$，根据定义 5.5.1 有

$$f^H(\boldsymbol{x};\ \boldsymbol{d}' + \boldsymbol{d}'') \leqslant f^H(\boldsymbol{x};\ \boldsymbol{d}') + f^C(\boldsymbol{x};\ \boldsymbol{d}'') \quad \forall \boldsymbol{d}', \boldsymbol{d}'' \in \mathscr{B}. \tag{5.6.1}$$

于是，存在 $\lambda > 0$ 对于 $\boldsymbol{d}' \in \operatorname{dom} f^H(\boldsymbol{x};\ \cdot)$ 有 $\boldsymbol{d} - \lambda\boldsymbol{d}' \in \operatorname{dom} f^C(\boldsymbol{x};\ \cdot)$. 由(5.6.1)得

$$f^H(\boldsymbol{x};\ \boldsymbol{d}) \leqslant \lambda f^H(\boldsymbol{x};\ \boldsymbol{d}') + f^C(\boldsymbol{x};\ \boldsymbol{d} - \lambda\boldsymbol{d}'), \tag{5.6.2}$$

因此 $\boldsymbol{d} \in \operatorname{dom} f^H(\boldsymbol{x};\ \cdot)$.

(2) 据(5.6.1)容易知道，$f^H(\boldsymbol{x};\ \cdot)$的上图象在 $f^C(\boldsymbol{x};\ \cdot)$的上图象中稠密，因此有

$$\liminf_{\boldsymbol{d}' \to \boldsymbol{d}} f^H(\boldsymbol{x};\ \boldsymbol{d}') \leqslant f^C(\boldsymbol{x};\ \boldsymbol{d}).$$

再由定理 5.4.1 知 $f^C(\boldsymbol{x};\ \boldsymbol{d})$关于 $\boldsymbol{d}$ 是下半连续的，故上式为等式.

(3) 已知 $\boldsymbol{d}' \in \operatorname{int}\operatorname{dom} f^C(\boldsymbol{x};\ \cdot) = \operatorname{dom} f^H(\boldsymbol{x};\ \cdot)$，在(5.6.2)中令 $\lambda \to 0$ 得 $f^C(\boldsymbol{x};\ \boldsymbol{d}) = f^H(\boldsymbol{x};\ \boldsymbol{d})$. 由于 $\operatorname{dom} f^H(\boldsymbol{x},\ \cdot)$ 是开集，由定理 5.5.1 可以推得 $f^C: \mathscr{B} \times \mathscr{B} \to R$, $(\boldsymbol{x}, \boldsymbol{d}) \mapsto f^C(\boldsymbol{x};\ \boldsymbol{d})$ 在$(\boldsymbol{x}, \boldsymbol{d}')$处是上半连续的.

(4) 由(3)和定理 5.4.1 可得.

(5) 根据(2)，我们有

$$\partial^C f(\boldsymbol{x}) = \{\boldsymbol{x}^* \in \mathscr{B}^* \mid \langle \boldsymbol{x}^*, \boldsymbol{d}\rangle \leqslant f^H(\boldsymbol{x};\ \boldsymbol{d}) \quad \forall \boldsymbol{d} \in \mathscr{B}\}.$$

由已知 $\operatorname{dom} f^H(\boldsymbol{x};\ \cdot) = \mathscr{B}$，只要证明对任意的 $\boldsymbol{x} \in \mathscr{B}$ 有

$$f^H(\boldsymbol{x};\ -\boldsymbol{d}) = (-f)^H(\boldsymbol{x};\ \boldsymbol{d}).$$

令 $a = f^H(\boldsymbol{x};\ -\boldsymbol{d})$，对任给定的 $\varepsilon > 0$，存在 α, β 和 $\eta > 0$，使得对任意的 $\boldsymbol{y} \in \operatorname{dom} f \cap (\boldsymbol{x} + \alpha B_{\mathscr{B}})$, $t \in (0, \beta)$ 和 $\boldsymbol{p} \in \boldsymbol{d} + \eta B_{\mathscr{B}}$，有

$$\frac{f(\boldsymbol{y} + t\boldsymbol{p}) - f(\boldsymbol{y})}{t} \leqslant a + \varepsilon.$$

取 $\alpha' \in (0, \alpha)$, $\beta' \in (0, \beta)$ 和 $\eta' \in (0, \beta)$，使得 $\alpha' + \beta'(\|\boldsymbol{d}\|$

$+\eta')\leqslant\alpha$，则对任意的 $\boldsymbol{x}'\in\mathrm{dom}f\cap(\boldsymbol{x}+\alpha'B_{\mathscr{B}})$，$\lambda\in f(\boldsymbol{x})+\alpha'B_R$，满足 $\lambda\geqslant-f(\boldsymbol{x}')$，$t\in(0,\beta')$ 和 $\boldsymbol{p}\in\boldsymbol{d}+\eta'B_{\mathscr{B}}$，有

$$\frac{-f(\boldsymbol{x}'+t\boldsymbol{p})-\lambda}{t}\leqslant\frac{f(\boldsymbol{x}')-f(\boldsymbol{x}'+t\boldsymbol{p})}{t}.$$

令 $\boldsymbol{y}=\boldsymbol{x}'+t\boldsymbol{p}$，上式为

$$\frac{-f(\boldsymbol{x}'+t\boldsymbol{p})-\lambda}{t}\leqslant\frac{f(\boldsymbol{y}-t\boldsymbol{p})-f(\boldsymbol{y})}{t}\leqslant\alpha+\varepsilon,$$

注意 $\boldsymbol{y}\in\mathrm{dom}f\cap(\boldsymbol{x}+\alpha B)$，则得 $(-f)^H(\boldsymbol{x};\boldsymbol{d})\leqslant f^H(\boldsymbol{x};-\boldsymbol{d})$. 在上式中交换 f 和 $-f$ 即可推出结论. ▯

定理 5.6.2　设 $f:\mathscr{B}\to R\cup\{+\infty\}$ 是实值函数，$\boldsymbol{x}^0\in\mathrm{dom}f$. 又设 $K\subset\mathscr{B}$ 是闭凸锥，并且对任意的 $\boldsymbol{x}\in\mathscr{B}$ 和任意的 $\boldsymbol{y}\in K$，有 $f(\boldsymbol{x}+\boldsymbol{y})\leqslant f(\boldsymbol{x})$.

(1) 对任意的 $\boldsymbol{d}\in K$，有 $f^C(\boldsymbol{x}^0;\boldsymbol{d})\leqslant 0$ 和 $\partial^C f(\boldsymbol{x}^0)\subset K^-$.

(2) 若 $\mathrm{int}K\neq\varnothing$，则对任意的 $\boldsymbol{d}\in\mathrm{int}K$ 有 $f^H(\boldsymbol{x}^0;\boldsymbol{d})\leqslant 0$.

证明　(1) 任给 $\varepsilon>0$ 和 $\beta>0$，对任意的 $\boldsymbol{x}\in\boldsymbol{x}^0+\varepsilon B$，$\alpha\in f(\boldsymbol{x}^0)+\varepsilon B_R$，满足 $\alpha\geqslant f(\boldsymbol{x})$，$t\in(0,\beta)$ 和 $\boldsymbol{d}\in K$，我们有

$$\frac{f(\boldsymbol{x}+t\boldsymbol{d})-\alpha}{t}\leqslant\frac{f(\boldsymbol{x}+t\boldsymbol{d})-f(\boldsymbol{x})}{t}\leqslant 0,$$

因此 $f^C(\boldsymbol{x};\boldsymbol{d})\leqslant 0$. 再由定义 5.4.2，可得 $\partial^C f(\boldsymbol{x})\subset K^-$.

(2) 因为 $\boldsymbol{d}\in\mathrm{int}K$，所以存在 $\bar\delta>0$，使得 $\boldsymbol{d}+\bar\delta B\subset\mathrm{int}K$. 据此，对任意的 $\varepsilon>0$ 和 $\beta>0$，$\boldsymbol{x}\in\boldsymbol{x}^0+\varepsilon B$，$\alpha\geqslant f(\boldsymbol{x})$，$t\in(0,\beta)$，以及 $\boldsymbol{d}'\in\boldsymbol{d}+\bar\delta B$，有

$$\frac{f(\boldsymbol{x}+t\boldsymbol{d}')-\alpha}{t}\leqslant\frac{f(\boldsymbol{x}+t\boldsymbol{d}')-f(\boldsymbol{x})}{t}.$$

根据定理 3.5.1，即得 $f^H(\boldsymbol{x}_0;\boldsymbol{d})\leqslant 0$. ▯

定理 5.6.3　设集合 $S\subset\mathscr{B}$ 非空，$\boldsymbol{x}^0\in S$.

(1) 若 $\boldsymbol{x}^*\in\mathscr{B}^*$ 使得 $\langle\boldsymbol{x}^*,\boldsymbol{x}^0\rangle=\max\limits_{\boldsymbol{x}\in S}\langle\boldsymbol{x}^*,\boldsymbol{x}\rangle$，则 $\boldsymbol{x}^*\in N_S^C(\boldsymbol{x}^0)$.

（2）若 $\mathscr{B}$ 是 Hilbert 空间，$y \notin \mathrm{cl}S$，$x \in \pi_{\mathrm{cl}S}(y)$ $\left(即 \|x - y\| = \min_{x' \in \mathrm{cl}S} \|y - x'\|\right)$，则 $y - x \in N_S^C(x)$.

证明 （1）作函数 $f:\mathscr{B} \to R$，$f(x) = \langle -x^*, x\rangle$. 因为 f 在 S 上的 x^0 处达到极小值，根据定理 5.3.2 和推论 4.4.4，得知 $\boldsymbol{0} \in \partial^C(f|_S)(x^0) \subset -x^* + N_S^C(x^0)$.

（2）作函数 $f:\mathscr{B} \to R$，$f(x) = \|x - y\|$. 由于 f 在 S 上的 x^0 处达到极小值，我们有

$$\boldsymbol{0} \in \partial^C\left(f|_S\right)(x) \subset \nabla f(x) + N_S^C(x)$$

$$= \frac{x - y}{\|x - y\|} + N_S^C(x),$$

因此 $y - x \in N_S^C(x)$. ▯

定理 5.6.4 设 $f:\mathscr{B} \to R \cup \{+\infty\}$ 是上半连续的正常函数，$c \in R$，$x^0 \in H(f, c) = \{x \in \mathscr{B} | f(x) \leqslant c\}$ 使得 $f(x^0) = c$.

（1）$T_{H(f,c)}^D(x^0) \subset \{d \in \mathscr{B} | f^D(x^0; d) \leqslant 0\}$.

（2）若存在 $\bar{d} \in \mathscr{B}$ 使得 $f^C(x^0; \bar{d}) < 0$，则 $\{d \in \mathscr{B} | f^C(x^0; d) \leqslant 0\} \subset T_{H(f,c)}^D(x^0)$.

证明 （1）设 $d \in T_{H(f,c)}^D(x^0)$，则存在序列 $\{d^k\}$，$d^k \to d$ 和 $\{t_k\}$，$t_k \to 0^+$ $(k \to \infty)$，使得 $x^0 + t_k d^k \in H(f, c)$ 对任何 $k = 1, 2, \cdots$ 成立. 由此，我们有 $f(x^0 + t_k d^k) \leqslant c (k = 1, 2, \cdots)$，故

$$\frac{f(x^0 + t_k d^k) - f(x^0)}{t_k} \leqslant 0,$$

从而 $f^D(x^0; d) \leqslant 0$.

（2）设 $a = -f^C(x_0; \bar{d})$，$a > 0$，则对任意的 $\varepsilon \in (0, a)$，有 $\alpha > 0$，$\beta > 0$，使对任意的 $x \in x^0 + \alpha B$ 和 $t \in (0, \beta)$，存在 $d \in \bar{d} + \varepsilon B$，使得

$$f(x + td) \leqslant f(x) + t(-a + \varepsilon).$$

因此，对任意的 $x \in B_{H(f,c)}(x^0, \alpha)$ 和 $t \in (0, \beta)$，存在 $\bar{d} \in \bar{d} + \varepsilon B$

使得 $f(x+td)\leqslant c$，这即 $x+td\in H(f,c)$，故 $\bar{d}\in T^C_{H(f,c)}(x^0)$。设 $d\in\mathscr{B}$ 满足 $f^C(x^0;d)\leqslant 0$。对 $\lambda\in(0,1)$，令 $d_\lambda=(1-\lambda)d+\lambda\bar{d}$，因为 $f^C(x^0;d)$ 关于 d 是凸的，故 $f^C(x^0;d_\lambda)<0$。由此，我们有 $d_\lambda\in T^C_{H(f,c)}(x^0)$，令 $\lambda\to 0$ 得 $d\in T^C_{H(f,c)}(x^0)$。▯

定理 5.6.5　设向量函数 $g:R^n\to R^m$ 是可微的，$f:R^m\to R\cup\{+\infty\}$ 是下半连续的正常函数，$x\in g^{-1}(\mathrm{dom}f)$。若

$$\mathrm{Im}Jg(x)-\mathrm{dom}\left(f^C(g(x);\ \cdot\ )\right)=R^m,$$

则对任意的 $d\in R^n$ 有

$$f^C\left(g(x);Jg(x)d\right)\geqslant(fg)^C(x;d)$$

和

$$\partial^C(fg)(x)\subset Jg(x)^*\partial^C f(g(x)).$$

证明　由推论 5.2.8 可推得。▯

定理 5.6.6　设 $f:R^n\to R\cup\{+\infty\}$ 是下半连续的正常函数，$S\subset R^n$ 是闭集。若 $x\in S\cap\mathrm{dom}f$，$\mathrm{dom}f^C(x;\ \cdot\ )-T^C_S(x)=R^n$，则对任意的 $d\in T^C_S(x)$，有 $(f|_S)^C(x;d)\leqslant f^C(x;d)$ 和

$$\partial^C(f|_S)(x)\subset\partial^C f(x)+N^C_S(x).$$

证明　根据定理 5.2.7 可以推得。▯

设 $\mathscr{D}$ 是 Banach 空间。

定理 5.6.7　设 $g:\mathscr{B}\times\mathscr{D}\to R\cup\{+\infty\}$ 是正常函数，$f:\mathscr{D}\to R$，

$$f(y)=\inf\{g(x,y)\,|\,x\in\mathscr{B}\}.$$

若 $x_y\in\mathscr{B}$ 使得 $g(x_y,y)=f(y)$，则对任意的 $p\in\mathscr{D}$ 有

$$f^D(y;d)\leqslant\inf\{g^D(x_y,y;d,p)\,|\,d\in\mathscr{B}\}.$$

证明　设 $d\in\mathscr{B}$，$p\in\mathscr{D}$，对任给的 $\epsilon>0$ 和 $\beta>0$，当 $d'\in d+\epsilon B_{\mathscr{B}}$，$p'\in p+\epsilon B_{\mathscr{D}}$ 和 $t\in(0,\beta)$ 时，有

$$\frac{f(y+tp')-f(y)}{t}\leqslant\frac{g(x_y+td',\ y+tp')-g(x_y,\ y)}{t}.$$

据此，得到

$$f^D(y;\ p)\leqslant g^D(x_y,\ y;\ d,\ p),\quad \forall(d,\ p)\in\mathscr{B}\times\mathscr{D},$$

因而结论成立. ▯

注 5.6.1 若 $f(y)=\sup\{g(x,\ y)\mid x\in\mathscr{B}\}$，则有

$$f^D(y;\ p)\geqslant\sup\{g^D(x_y,\ y;\ d,\ p)\mid d\in\mathscr{B}\}.$$

定理 5.6.8 设 $f_i:\mathscr{R}^n\to R\cap\{+\infty\}(i=1,\cdots,n)$ 是下半连续的正常函数，$f(x)=\max\{f_i(x)\mid i=1,\cdots,m\}$. 又设

$$x^0\in\bigcap_{i=1}^{m}\operatorname{int}\operatorname{dom}f_i,\ J(x^0)=\{i=1,\ \cdots,\ m\mid f_i(x^0)=f(x^0)\}.$$

若对任何 $i\in J(x^0)$ 和任何 $d^i\in\mathscr{B}$ 存在 $\bar{d}$，使得 $\bar{d}-d^i\in\operatorname{dom}f_i^C(x^0;\ \cdot)$，则对任意的 $d\in\mathscr{B}$ 有 $f^C(x^0;\ d)\leqslant\max\limits_{i\in J(x^0)}f_i^C(x^0;\ d)$，并且

$$\partial^C f(x_0)\subset\overline{\mathrm{co}}\bigcup_{i\in J(x^0)}\partial^C f_i(x^0).$$

证明 已知 $x^0\in\bigcap\limits_{i=1}^{m}\operatorname{int}\operatorname{dom}f_i$，当 $i\notin J(x^0)$，则 $(x^0,\ f(x^0))\in\operatorname{int}(\operatorname{epi}f_i)$，因此切锥 $T^C_{\operatorname{epi}f_i}(x^0,\ f(x^0))=\mathscr{B}\times R$. 再由已设，对所有的 $i=1,\ \cdots,\ m$，存在 $\bar{d}$ 使得 $\bar{d}-d^i\in\operatorname{dom}f_i^C(x^0;\ \cdot)$，则对任意的 $(d^i,\ \lambda_i)\in\mathscr{B}\times R(k=1,\ \cdots,\ m)$ 存在 $(d,\ \bar{\lambda})$，使得

$$(\bar{d}-d^i,\ \bar{\lambda}-\lambda_i)\in T^C_{\operatorname{epi}f_i}(x^0,\ f(x^0)),\ i=1,\ \cdots,\ m,$$

其中 $\bar{\lambda}=\max\limits_{i\in J(x^0)}\{f_i^C(x^0;\ \bar{d}-d^i)+\lambda_i\}$. 根据推论 5.2.10，因为 $\operatorname{epi}f=\bigcap\limits_{i=1}^{m}\operatorname{epi}f_i$，所以有

$$\bigcap_{i\in J(x^0)}^{m} T^C_{\mathrm{epi}f_i}\left(x^0,\ f(x^0)\right)$$

$$=\bigcap_{i=1}^{m} T^C_{\mathrm{epi}f_i}\left(x^0,\ f(x^0)\right)\subset T^C_{\mathrm{epi}f}\left(x^0,\ f(x^0)\right),$$

于是得知结论成立. □

推论 5.6.9　设 $f_i:R^n\to R\cup\{+\infty\}(i=1,\cdots,m)$ 是下半连续的正常函数. 若 $f_i(i=1,\cdots,m)$ 在 $x\in R^n$ 处是 Lipschitz 的,则对任意的 $d\in R^n$ 有

$$f^C(x;\ d)\leqslant \max_{i\in J(x^0)} f_i^C(x;\ d),$$

$$\partial^C f(x)\subset \overline{\mathrm{co}}\bigcup_{i\in J(x^0)}\partial^C f_i(x).$$

证明　由定理 5.6.8 可推得. □

定理 5.6.10　设 $f,\ g:\mathscr{B}\to R\cup\{+\infty\}$ 是正常函数, $x\in \mathrm{dom}f\cap\mathrm{dom}g$. 若 g 在 x 处是正常 H-可切导的,并且 $d\in\mathrm{dom}f^C(x;\ \cdot)\cap\mathrm{dom}g^H(x;\ \cdot)\neq\varnothing$,则

$$(f+g)^C(x;\ d)\leqslant f^C(x;\ d)+g^C(x;\ d),$$

$$\partial^C(f+g)(x)\subset\partial^C f(x)+\partial^C g(x).$$

证明　对任意的 $d\in\mathrm{dom}f^C(x;\ \cdot)\cap\mathrm{dom}g^H(x;\ \cdot)$, 有

$$(f+g)^C(x;\ d)\leqslant f^C(x;\ d)+g^H(x;\ d),\qquad(5.6.3)$$

设 $d^0\in\mathrm{dom}f^C(x;\ \cdot)\cap\mathrm{dom}g^C(x;\ \cdot)$ 和 $\lambda\in(0,\ 1)$, 由定理 5.6.1 有 $\mathrm{dom}g^H(x;\ \cdot)=\mathrm{int\,dom}f^C(x;\ \cdot)$, 所以

$$(1-\lambda)d^0+\lambda d\in\mathrm{dom}f^C(x;\ \cdot)\cap\mathrm{dom}g^H(x;\ \cdot).$$

依据(5.6.3),得到

$$(f+g)^C\left(x;\ (1-\lambda)d^0+\lambda d\right)$$

$$\leqslant(1-\lambda)f^C(x;\ d)+\lambda f^C(x;\ d)+g^H(x,\ (1-\lambda)d^0+\lambda d).$$

另外,由(5.6.1)可得

$$g^H\big(x;(1-\lambda)d^0+\lambda d\big)\leqslant(1-\lambda)g^C(x;d^0)+\lambda g^H(x;d).$$

在以上两式中令 $\lambda\to 0$,由(5.6.3)得第1个结论,再由推论5.4.4得第2个结论. □

最后,给出有关复合映射的定理如下.

定理 5.6.11 设映射 $\varphi:\mathscr{B}\to\mathscr{D}$ 是严格可微的,$g:\mathscr{D}\to R\cup\{+\infty\}$ 是正常函数,$f=g(\varphi(x)):\mathscr{B}\to R\cup\{+\infty\}$,并且 g 在 $\varphi(x)$ 处是有限的且具有有限 H-切导数.若

$$\mathrm{dom}D_S\varphi(x)\cap\mathrm{int}\{d\in\mathscr{B}\,|\,g^C(\varphi(x);d)<+\infty\}\neq\varnothing,$$

则

$$\partial^C f(x)\subset D_S\varphi(x)^*\partial^C g\big(\varphi(x)\big).$$

再若 g 在 x 处是正则的,则上式成为等式.

证明 设 $y=\varphi(x)$,作映射 $h:\mathscr{B}\times\mathscr{D}\to R\cup\{+\infty\}$,

$$h(x',y')=\begin{cases}f(x'), & y'=\varphi(x');\\ +\infty, & \text{其他}.\end{cases}$$

令 $h=f_1+f_2$,其中 f_1 是 φ 图象的指示函数,$f_2(x',y')=g(y')$,$F=D_S\varphi(x)$,$G=\mathrm{graph}\varphi$.我们证明下面的结论成立:

$$h^C(x,y;d,p)=\begin{cases}f^C(x;d), & p=F(d);\\ +\infty, & \text{其他}.\end{cases}$$

事实上,有

$$f^C(x,y;d,p)$$

$$=\lim_{\substack{(x',y',\alpha)\downarrow(x,y)\\ t\to 0^+}}\sup_{}\inf_{\substack{d'\to d\\ p'\to p}}\frac{h(x'+td',y'+tp')-\alpha}{t},$$

其中 $\alpha\geqslant h(x',y')=f(x')$,$y'=\varphi(x')$,并且

$$h(x'+td',\ y'+tp')=\begin{cases} f(x'+td'), & p'=\dfrac{\varphi(x'+td')-\varphi(x')}{t}; \\ +\infty, & \text{其他}. \end{cases}$$

因为 φ 在 x 处是严格可微的，我们有

$$\lim_{\substack{x'\to x\ d'\to d \\ t\to 0^+}} \sup \frac{\varphi(x'+td')-\varphi(x')}{t}=F(d).$$

因此，当 $p=F(d)$ 时，有

$$\lim_{\substack{(x',a)\downarrow_f x\ d'\to d \\ t\to 0^+}} \sup\inf \frac{f(x'+td')-a}{t}=f^C(x;d),$$

并且当 $p\neq F(d)$ 时有 $f^C(x,y;d,p)=+\infty$. 由定义 5.4.2 推知

$$\begin{aligned}\partial^C h(x,y) &= \{(x^*,y^*)\,|\,h^C(x,y;d,p)\geqslant\langle x^*,d\rangle \\ &\quad +\langle y^*,p\rangle\quad \forall d\in\mathscr{B},\forall p\in\mathscr{D}\} \\ &= \{(x^*,y^*)\,|\,f^C(x,d)\geqslant\langle x^*,d\rangle \\ &\quad +\langle y^*,F(d)\rangle\quad\forall d\in\mathscr{B}\} \\ &= \{(x^*,y^*)\,|\,f^C(x,d)\geqslant\langle x^*+F^*(y^*),d\rangle\quad\forall d\in\mathscr{B}\} \\ &= \{(x^*,y^*)\,|\,x^*+F^*(y^*)\in\partial^C f(x)\},\end{aligned}$$

因此得

$$\partial^C f(x)=\{x^*\,|\,(x^*,0)\in\partial^C h(x,y)\}. \tag{5.6.4}$$

从 φ 是严格可微的，我们有 $T_G^C(x,y)=T_G^b(x,y)=\text{graph}\varphi$，故 f_1 在 (x,y) 处是正则的，并且有

$$f_1^C(x,y;d,p)=\begin{cases}0, & p=F(d);\\ +\infty, & \text{其他},\end{cases}$$

$$\partial^C f_1(x,y)=N_G^C(x,y)=\{(x^*,y^*)\,|\,x^*=-F^*(y^*)\}. \tag{5.6.5}$$

显然，f_2 在 $(\boldsymbol{x}, \boldsymbol{y})$ 处是有限正常 H-可切导的，从而有

$$f_2^C(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}, \boldsymbol{p}) = g^C(\boldsymbol{y}; \boldsymbol{p}),$$

$$\partial^C f_2(\boldsymbol{x}, \boldsymbol{y}) = \{(\boldsymbol{0}, \boldsymbol{y}^*) \mid \boldsymbol{y}^* \in \partial^C g(\boldsymbol{y})\}. \qquad (5.6.6)$$

由已知得

$$\begin{aligned} &\operatorname{dom} f_1^C(\boldsymbol{x}, \boldsymbol{y}; \cdot, \cdot) \cap \operatorname{dom} f_2^H(\boldsymbol{x}, \boldsymbol{y}; \cdot, \cdot) \\ &= \{(\boldsymbol{d}, \boldsymbol{F}(\boldsymbol{d})) \mid \boldsymbol{F}(\boldsymbol{d}) \in \operatorname{int}\{\boldsymbol{p} \mid g^C(\boldsymbol{y}; \boldsymbol{p}) < +\infty\}\}, \end{aligned}$$

利用定理 5.6.9 有

$$\partial^C h(\boldsymbol{x}, \boldsymbol{y}) \subset \partial^C f_1(\boldsymbol{x}, \boldsymbol{y}) + \partial^C f_2(\boldsymbol{x}, \boldsymbol{y}).$$

根据(5.6.5)、(5.6.6)、(5.6.4)和上式即可得到结论. 再若 g 在 $\boldsymbol{x}$ 处是正则的，从上式不难推出结论对等式情形成立. ▯

第 6 章　锥凸集值映射的锥次微分

从这一章开始，我们研究集值映射的微分问题. 集值映射的次微分是现代非光滑分析理论的重要组成部分.

本章考虑在由尖闭凸锥确定偏序的局部凸线性拓扑空间，对锥凸集值映射引进它的锥次微分和锥弱次微分概念. 我们指出，本章和下一章介绍的关于集值映射的次微分都是在有效性意义下给出的. 可以看到，这些次微分仍然具有一系列重要性质. 在本章最后，定义了集值映射的锥方向导数和锥弱方向导数.

§6.1　锥次微分和锥弱次微分

设 $\mathscr{X}$ 和 $\mathscr{Y}$ 是局部凸线性拓扑空间，$\mathscr{X}^*$ 和 $\mathscr{Y}^*$ 分别是 $\mathscr{X}$ 和 $\mathscr{Y}$ 的对偶空间，$\langle \boldsymbol{x}^*, \boldsymbol{x}\rangle$ 表示 $\mathscr{X}$ 中连续线性泛函 $\boldsymbol{x}^*$ 在 $\boldsymbol{x} \in \mathscr{X}$ 处的值，$\langle \boldsymbol{y}^*, \boldsymbol{y}\rangle$ 表示连续线性泛函 $\boldsymbol{y}^*$ 在 $\boldsymbol{y} \in \mathscr{Y}$ 处的值. 除特别说明外，以下均设 $K \subset \mathscr{Y}$ 是尖闭凸锥，$\mathscr{Y}$ 中的序由 K 确定. 设 $\boldsymbol{x} \in \mathscr{X}$，$\psi$: $\mathscr{X} \to 2^{\mathscr{Y}}$，$\boldsymbol{x} \mapsto \psi(\boldsymbol{x})$ 是集值映射.

先引进集值映射的锥凸性，它是函数凸性的推广.

定义 6.1.1　设 $S \subset \mathscr{X}$ 是非空凸集，集合 $\psi(\boldsymbol{x}) \subset \mathscr{Y}$ $(\boldsymbol{x} \in S)$，ψ: $S \to 2^{\mathscr{Y}}$，$\boldsymbol{x} \mapsto \psi(\boldsymbol{x})$ 是集值映射，$\mathrm{int}K \neq \varnothing$.

(1) 若对任意的 $\lambda \in (0, 1)$ 有

$$\begin{aligned} &\lambda\psi(\boldsymbol{x}^1) + (1-\lambda)\psi(\boldsymbol{x}^2) \\ &\subset \psi\left(\lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2\right) + K \quad \forall \boldsymbol{x}^1, \boldsymbol{x}^2 \in S, \end{aligned} \tag{6.1.1}$$

则称 ψ 是 S 上的 K-凸集值映射，或 ψ 在 S 上是 K-凸的. 在(6.1.1)中令 $\boldsymbol{x}^1=\boldsymbol{x}^0$ 和 $\boldsymbol{x}^2=\boldsymbol{x}$，则称 ψ 在点 $\boldsymbol{x}^0$ 处(相对于 S)是 K-凸的.

(2) 若对任意的 $\lambda\in(0,1)$ 有

$$\psi(\boldsymbol{x}^1)+(1-\lambda)\psi(\boldsymbol{x}^2)$$

$$\subset\psi\big(\lambda\boldsymbol{x}^1+(1-\lambda)\boldsymbol{x}^2\big)+\text{int}K\quad\forall\boldsymbol{x}^1,\boldsymbol{x}^2\in S,\tag{6.1.2}$$

则称 ψ 是 S 上的 K-严格凸集值映射，或 ψ 在集合 S 上是 K-严格凸的. 在(6.1.2)中令 $\boldsymbol{x}^1=\boldsymbol{x}^0$ 和 $\boldsymbol{x}^2=\boldsymbol{x}$，则称 ψ 在点 $\boldsymbol{x}^0$ 处(相对于 S)是 K-严格凸的.

注 6.1.1 在定义 6.1.1 的(1)中，当集值映射退化为实值函数 $\psi=f$: $S\to R$，则它在 S 上是 R_+-凸的，即 f 是 S 上的凸函数.

下面介绍集合的锥有效性.

定义 6.1.2 设集合 $Y\subset\mathscr{Y}$ 非空，$\text{int}K\neq\varnothing$.

(1) 设 $\bar{\boldsymbol{y}}\in Y$，若不存在 $\boldsymbol{y}\in Y$ 使得

$$\bar{\boldsymbol{y}}-\boldsymbol{y}\in K\backslash\{\boldsymbol{0}\},$$

则称 $\bar{\boldsymbol{y}}$ 是集合 Y 的 K-有效点. Y 的所有 K-有效点组成的集合记作 $\mathscr{E}(Y,K)$.

(2) 设 $\bar{\boldsymbol{y}}\in Y$，若不存在 $\boldsymbol{y}\in Y$ 使得

$$\bar{\boldsymbol{y}}-\boldsymbol{y}\in\text{int}K,$$

则称 $\bar{\boldsymbol{y}}$ 是集合 Y 的 K-弱有效点. Y 的所有 K-弱有效点组成的集合记作 $\mathscr{E}_w(Y,K)$.

由定义 6.1.2，显然有

$$\mathscr{E}(Y,K)\subset\mathscr{E}_w(Y,K).\tag{6.1.3}$$

下面给出集值映射的锥次微分和锥弱次微分概念.

定义 6.1.3 设集合 $S\subset\mathscr{X}$ 和 $\psi(\boldsymbol{x})\subset\mathscr{Y}(\boldsymbol{x}\in S)$ 非空. ψ: $S\to2^{\mathscr{Y}}$, $\boldsymbol{x}\mapsto\psi(\boldsymbol{x})$ 是集值映射，点 $\boldsymbol{x}^0\in S$, $\boldsymbol{y}^0\in\psi(\boldsymbol{x}^0)$，并且 $\boldsymbol{x}^*\in\mathscr{X}^*$. 又 $\text{int}K\neq\varnothing$，给定向量 $\boldsymbol{p}\in\text{int}K$.

(1) 若不存在 $y \in \bigcup_{x \in S}[\psi(x) - \langle x^*, x\rangle p]$ 使得

$$y^0 - \langle x^*, x^0\rangle p - y \in K\setminus\{\theta\},$$

则称 x^* 是集值映射 ψ 在点(x^0, y^0)处关于向量 p 的 K-次梯度，ψ 在(x^0, y^0)处关于 p 的所有 K-次梯度组成的集合称为 ψ 在点(x^0, y^0) 处关于 p 的 K-次微分，记作$\partial\psi(x^0, y^0)_p$. 若 $\partial\psi(x^0, y^0)_p \neq \varnothing$，则称 ψ 在点(x^0, y^0)处关于 p 是 K-次可微的. 若集合

$$\partial\psi(x^0)_p = \{x^* \in \partial\psi(x^0, y^0)_p \mid y^0 \in \psi(x^0)\} \neq \varnothing, \tag{6.1.4}$$

则称集值映射 ψ 在点 x^0 处关于向量 p 是 K-次可微的，这时称集合$\partial\psi(x^0)_p$ 是 ψ 在点 x^0 处关于 p 的 K-次微分.

(2) 若不存在 $y \in \bigcup_{x \in S}[\psi(x) - \langle x^*, x\rangle p]$ 使得

$$y^0 - \langle x^*, x^0\rangle p - y \in \operatorname{int}K,$$

则称 x^* 是集值映射 ψ 在点(x^0, y^0)处关于向量 p 的 K-弱次梯度，ψ 在(x^0, y^0)处关于 p 的所有 K-弱次梯度组成的集合称为 ψ 在点(x^0, y^0)处关于 p 的 K-弱次微分. 记作 $\partial_w\psi(x^0, y^0)_p$. 若 $\partial_w\psi(x^0, y^0)_p \neq \varnothing$，则称 ψ 在点(x^0, y^0)处关于 p 是 K-弱次可微的，若集合

$$\partial_w\psi(x^0)_p = \{x^* \in \partial_w\psi(x^0, y^0)_p \mid y^0 \in \psi(x^0)\} \neq \varnothing, \tag{6.1.5}$$

则称集值映射 ψ 在点 x^0 处关于向量 p 是 K-弱次可微的，这时称集合 $\partial_w\psi(x^0)_p$ 是 ψ 在点 x^0 处关于 p 的 K-弱次微分.

按定义 6.1.2 和定义 6.1.3，则有

$$\begin{aligned}\partial\psi(x^0, y^0)_p = \Big\{x^* \in \mathscr{L}^* \mid\, & y^0 - \langle x^*, x^0\rangle p \\ & \in \mathscr{E}\Big\{\bigcup_{x \in S}[\psi(x) - \langle x^*, x\rangle p], K\Big\}\Big\},\end{aligned} \tag{6.1.6}$$

$$\partial_w\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)_p = \Big\{\boldsymbol{x}^* \in \mathscr{X}^* \,\Big|\, \boldsymbol{y}^0 - \langle \boldsymbol{x}^*, \boldsymbol{x}^0\rangle \boldsymbol{p} \in \mathscr{E}_w\Big(\bigcup_{\boldsymbol{x}\in S}\big[\psi(\boldsymbol{x}) - \langle \boldsymbol{x}^*, \boldsymbol{x}\rangle \boldsymbol{p}\big], K\Big)\Big\}. \tag{6.1.7}$$

由定义 6.1.3，显然有

$$\partial\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)_p \subset \partial_w\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)_p,$$

$$\partial\ \psi(\boldsymbol{x}^0)_p \subset \partial_w\psi(\boldsymbol{x}^0)_p. \tag{6.1.8}$$

注 6.1.2 在定义 6.1.3 的(1)中，当集值映射退化为实值函数 $\psi = f\colon S \to R$，并且 $K = R_+$ 和 $p = 1$ 时，有

$$f(\boldsymbol{x}^0) - \langle \boldsymbol{x}^*, \boldsymbol{x}^0\rangle \in \mathscr{E}\Big(\bigcup_{\boldsymbol{x}\in S}\big[f(\boldsymbol{x}) - \langle \boldsymbol{x}^*, \boldsymbol{x}\rangle\big], R_+\Big),$$

按定义 6.1.2 的(1)，即不存在 $\boldsymbol{x} \in S$，使得

$$f(\boldsymbol{x}) - \langle \boldsymbol{x}^*, \boldsymbol{x}\rangle \leqslant f(\boldsymbol{x}^0) - \langle \boldsymbol{x}^*, \boldsymbol{x}^0\rangle,$$

或即

$$\langle \boldsymbol{x}^*, \boldsymbol{x} - \boldsymbol{x}^0\rangle \leqslant f(\boldsymbol{x}) - f(\boldsymbol{x}^0) \quad \forall \boldsymbol{x} \in S.$$

由此可见，这时 f 在点 $\boldsymbol{x}^0$ 处关于 $p = 1$ 的 R_+—次微分，即是第 3 章中关于凸函数的次微分. 因此，定义 6.1.3 是凸函数的次梯度和次微分的一种推广.

关于锥次微分和锥弱次微分，有以下基本结果.

定理 6.1.1 设集合 $S \subset \mathscr{X}$ 非空，$\psi\colon S \to 2^{\mathscr{Y}}$ 是集值映射，$\boldsymbol{x} \in S$，$\boldsymbol{y} \in \psi(\boldsymbol{x})$，$\boldsymbol{p} \in \operatorname{int}K$.

(1) 若 $\boldsymbol{x}^* \in \partial\psi(\boldsymbol{x}, \boldsymbol{y})_p$，则 $\boldsymbol{y} \in \mathscr{E}\big(\psi(\boldsymbol{x}), K\big)$；若 $\boldsymbol{x}^* \in \partial_w\psi(\boldsymbol{x}, \boldsymbol{y})_p$，则 $\boldsymbol{y} \in \mathscr{E}_w\big(\psi(\boldsymbol{x}), K\big)$.

(2) 若 $\partial\psi(\boldsymbol{x})_p \neq \varnothing$，则 $\mathscr{E}\big(\psi(\boldsymbol{x}), K\big) \neq \varnothing$；若 $\partial_w\psi(\boldsymbol{x})_p \neq \varnothing$，则 $\mathscr{E}_w\big(\psi(\boldsymbol{x}), K\big) \neq \varnothing$.

(3) 若 $\boldsymbol{y} \notin \mathscr{E}\big(\psi(\boldsymbol{x}), K\big)$，则 $\partial\,\psi(\boldsymbol{x}, \boldsymbol{y})_p = \varnothing$；若 $\boldsymbol{y} \notin$

$\mathscr{E}_w\big(\psi(\boldsymbol{x}),\ K\big)$，则 $\partial_w\psi(\boldsymbol{x},\ \boldsymbol{y})_{\boldsymbol{p}}=\varnothing$.

证明　(1) 反之，假设 $\boldsymbol{y}\notin\mathscr{E}\big(\psi(\boldsymbol{x}),\ K\big)$，由定义 6.1.2 可知存在 $\boldsymbol{y}'\neq\boldsymbol{y},\ \boldsymbol{y}'\in\psi(\boldsymbol{x})$，使得 $\boldsymbol{y}-\boldsymbol{y}'\in K\backslash\{\boldsymbol{0}\}$，由此有

$$\boldsymbol{y}-\langle\boldsymbol{x}^*,\ \boldsymbol{x}\rangle\boldsymbol{p}-\big(\boldsymbol{y}'-\langle\boldsymbol{x}^*,\ \boldsymbol{x}\rangle\boldsymbol{p}\big)\in K\backslash\{\boldsymbol{0}\}.$$

根据定义 6.1.3 的(1)得知 $\boldsymbol{x}^*\notin\partial\psi(\boldsymbol{x},\ \boldsymbol{y})_{\boldsymbol{p}}$，导致矛盾. 类似地可证第 2 个结论.

(2)和(3)的第 1 个结论由(1)的第 1 个结论可推得，第 2 个结论由(1)的第 2 个结论可推得. ▯

定理 6.1.2　设集合 $S\subset\mathscr{X}$ 非空，$\psi:\ S\to2^{\mathscr{Y}}$ 是集值映射，$\boldsymbol{x}\in S,\ \boldsymbol{y}\in\psi(\boldsymbol{x}),\ \boldsymbol{p}\in\mathrm{int}K,\ \delta>0$.

(1) 若 $\boldsymbol{x}^*\in\partial\psi(\boldsymbol{x},\ \boldsymbol{y})_{\boldsymbol{p}}$，则 $\boldsymbol{x}^*/\delta\in\partial\psi(\boldsymbol{x},\ \boldsymbol{y})_{\delta\boldsymbol{p}}$.

(2) 若 $\boldsymbol{x}^*\in\partial_w\psi(\boldsymbol{x},\ \boldsymbol{y})_{\boldsymbol{p}}$，则 $\boldsymbol{x}^*/\delta\in\partial_w\psi(\boldsymbol{x},\ \boldsymbol{y})_{\delta\boldsymbol{p}}$.

证明　由定义 6.1.3 不难推得. ▯

定理 6.1.3　设集合 $S\subset\mathscr{X}$ 非空，$\boldsymbol{x}\in S,\ \boldsymbol{y}\in\psi(\boldsymbol{x}),\ \boldsymbol{p}\in\mathrm{int}K$.

(1) $\boldsymbol{0}\in\partial\psi(\boldsymbol{x},\ \boldsymbol{y})_{\boldsymbol{p}}$ 当且仅当 $\boldsymbol{y}\in\mathscr{E}\Big(\bigcup\limits_{\boldsymbol{x}\in S}\psi(\boldsymbol{x}),\ K\Big)$.

(2) $\boldsymbol{0}\in\partial_w\psi(\boldsymbol{x},\ \boldsymbol{y})_{\boldsymbol{p}}$ 当且仅当 $\boldsymbol{y}\in\mathscr{E}_w\Big(\bigcup\limits_{\boldsymbol{x}\in S}\psi(\boldsymbol{x}),\ K\Big)$.

证明　由定义 6.1.2 容易推得. ▯

定理 6.1.4　设集合 $S\subset\mathscr{X}$ 非空，$\psi:\ S\to2^{\mathscr{Y}}$ 是集值映射，$\psi(\boldsymbol{x}')\neq\varnothing(\forall\boldsymbol{x}'\in S)$. 若 ψ 在 S 上是 K-严格凸的，$\boldsymbol{x}\in S,\ \boldsymbol{y}\in\psi(\boldsymbol{x})$，则

$$\partial\psi(\boldsymbol{x},\ \boldsymbol{y})_{\boldsymbol{p}}=\partial_w\psi(\boldsymbol{x},\ \boldsymbol{y})_{\boldsymbol{p}}\quad\partial\ \psi(\boldsymbol{x})_{\boldsymbol{p}}=\partial_w\psi(\boldsymbol{x})_{\boldsymbol{p}}.$$

证明　记

$$Y(\boldsymbol{x}^*)=\bigcup_{\boldsymbol{x}\in S}\big[\psi(\boldsymbol{x})-\langle\boldsymbol{x}^*,\ \boldsymbol{x}\rangle\boldsymbol{p}\big],\tag{6.1.9}$$

由(6.1.3)我们有

$$\mathscr{E}\big(Y(\boldsymbol{x}^*),\ K\big)\subset\mathscr{E}_w\big(Y(\boldsymbol{x}^*),\ K\big).\tag{6.1.10}$$

下面证明也有

$$\mathscr{E}_w(Y(\boldsymbol{x}^*),\ K) \subset \mathscr{E}(Y(\boldsymbol{x}^*),\ K). \tag{6.1.11}$$

事实上,反之假设有 $\boldsymbol{y} \in \mathscr{E}_w(Y(\boldsymbol{x}^*),\ K)$ 和 $\boldsymbol{y} \notin \mathscr{E}(Y(\boldsymbol{x}^*),\ K)$,从后者按定义 6.1.2 的(1) 可知,存在 $\boldsymbol{y}' \in Y(\boldsymbol{x}^*)$ 使得

$$\boldsymbol{y} - \boldsymbol{y}' \in K\backslash\{\boldsymbol{0}\}. \tag{6.1.12}$$

由 $\boldsymbol{y} \in \mathscr{E}_w(Y(\boldsymbol{x}^*),\ K) \subset Y(\boldsymbol{x}^*)$ 和 $\boldsymbol{y}' \in Y(\boldsymbol{x}^*)$,根据(6.1.9) 知,存在 $\boldsymbol{x}^1,\ \boldsymbol{x}^2 \in S,\ \boldsymbol{y}^1 \in \psi(\boldsymbol{x}^1),\ \boldsymbol{y}^2 \in \psi(\boldsymbol{x}^2)$,使 $\boldsymbol{y} = \boldsymbol{y}^1 - \langle \boldsymbol{x}^*,\ \boldsymbol{x}^1\rangle \boldsymbol{p}$, $\boldsymbol{y}' = \boldsymbol{y}^2 - \langle \boldsymbol{x}^*,\ \boldsymbol{x}^2\rangle \boldsymbol{p}$. 因此,对任意的 $\lambda \in (0,\ 1)$ 有

$$\begin{aligned} &\lambda\boldsymbol{y} + (1-\lambda)\boldsymbol{y}' \\ &= \lambda\boldsymbol{y}^1 + (1-\lambda)\boldsymbol{y}^2 - \left\langle \boldsymbol{x}^*,\ \lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 \right\rangle \boldsymbol{p}, \end{aligned} \tag{6.1.13}$$

由于 ψ 在 S 上是 K-严格凸的,按定义 6.1.1 的(2)得知

$$\lambda\boldsymbol{y}^1 + (1-\lambda)\boldsymbol{y}^2 \in \psi(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2) + \mathrm{int}K.$$

再由(6.1.13)和(6.1.9)得到

$$\lambda\boldsymbol{y} + (1-\lambda)\boldsymbol{y}' \in Y(\boldsymbol{x}^*) + \mathrm{int}K.$$

据此,存在 $\boldsymbol{y}'' \in Y(\boldsymbol{x}^*)$ 和 $\boldsymbol{v} \in \mathrm{int}K$,使 $\boldsymbol{y}'' + \boldsymbol{v} = \lambda\boldsymbol{y} + (1-\lambda)\boldsymbol{y}'$,故有 $\boldsymbol{y} - \boldsymbol{y}'' = (1-\lambda)(\boldsymbol{y} - \boldsymbol{y}') + \boldsymbol{v}$. 注意到(6.1.12) 和 K 是凸锥,可知

$$\boldsymbol{y} - \boldsymbol{y}'' = (1-\lambda)(\boldsymbol{y} - \boldsymbol{y}') + \boldsymbol{v} \in K\backslash\{\boldsymbol{0}\} + \mathrm{int}K \subset \mathrm{int}K.$$

于是,由定义 6.1.2 的(2)得到 $\boldsymbol{y} \notin \mathscr{E}_w(Y(\boldsymbol{x}^*),\ K)$,这导致与假设矛盾.

由(6.1.10)和(6.1.11)我们有 $\mathscr{E}(Y(\boldsymbol{x}^*),\ K) = \mathscr{E}_w(Y(\boldsymbol{x}^*),\ K)$,由此根据定义 6.1.3 即得结论. ▯

§6.2 存在性定理

本节研究集值映射的锥次微分和锥弱次微分的存在性.

设集合 $S \subset \mathscr{X}$，$\psi: S \to 2^{\mathscr{Y}}$，$\boldsymbol{x} \mapsto \psi(\boldsymbol{x})$ 是集值映射，$K \subset \mathscr{Y}$ 是内部非空的尖闭凸锥.

定义 6.2.1　设集合 $S \subset \mathscr{X}$ 非空，点 $\boldsymbol{x}^0 \in S$，$\psi: S \to 2^{\mathscr{Y}}$ 是集值映射. 若存在点 $\boldsymbol{a} \in \mathscr{Y}$ 和集合 $T \subset \mathscr{Y}$，使得 $\psi(\boldsymbol{x}^0) \subset T \subset \boldsymbol{a} - K$ 和 $\psi^{-1}(T) = \{\boldsymbol{x} \in S \mid \psi(\boldsymbol{x}) \subset T\}$ 是开集，则称 ψ 在点 $\boldsymbol{x}^0$ 处是 K-有界逆开的.

记集值映射 ψ 在 S 上的 K-上图象为

$$K\text{-epi}\psi = \{(\boldsymbol{x}, \boldsymbol{y}) \in \mathscr{X} \times \mathscr{Y} \mid \boldsymbol{x} \in S,\ \boldsymbol{y} \in \psi(\boldsymbol{x}) + K\}. \tag{6.2.1}$$

引理 6.2.1　设集合 $S \subset \mathscr{X}$ 非空，$\boldsymbol{x} \in S$. 若集值映射 $\psi: S \to 2^{\mathscr{Y}}$ 在 $\boldsymbol{x}$ 处是 K- 有界逆开的，则 $\mathrm{int}(K\text{-epi}\psi) \neq \varnothing$.

证明　因 ψ 在 $\boldsymbol{x}$ 处是 K-有界逆开的，按定义 6.2.1 可知，存在 $\boldsymbol{a} \in \mathscr{Y}$ 和 $T \subset \mathscr{Y}$，使得对任意的 $\boldsymbol{x} \in \psi^{-1}(T)$，取 $\boldsymbol{y}_x \in \psi(\boldsymbol{x})$，则存在 $\boldsymbol{p}_x \in K$ 使 $\boldsymbol{y}_x = \boldsymbol{a} - \boldsymbol{p}_x$. 设 $\overline{\boldsymbol{y}} = \boldsymbol{a} + \boldsymbol{p}$，则 $\overline{\boldsymbol{y}} - \boldsymbol{a} = \boldsymbol{p} \in \mathrm{int}K$，于是存在 $\boldsymbol{0}$ 点的邻域 $U(\boldsymbol{0}) \subset \mathscr{Y}$ 有

$$\boldsymbol{y} + \overline{\boldsymbol{y}} - \boldsymbol{a} \in K \quad \forall \boldsymbol{y} \in U(\boldsymbol{0}).$$

因为 K 是凸锥，所以从上式可得

$$\boldsymbol{y} + \overline{\boldsymbol{y}} - \boldsymbol{y}_x = \boldsymbol{y} + \overline{\boldsymbol{y}} - \boldsymbol{a} + \boldsymbol{p}_x \subset K$$

$$\forall \boldsymbol{x} \in \psi^{-1}(T),\ \forall \boldsymbol{y} \in U(\boldsymbol{0}),$$

或即

$$\boldsymbol{y} + \overline{\boldsymbol{y}} \subset \boldsymbol{y}_x + K \subset \psi(\boldsymbol{x}) + K$$

$$\forall \boldsymbol{x} \in \psi^{-1}(T),\ \forall \boldsymbol{y} \in U(\boldsymbol{0}).$$

据此，由(6.2.1)得

$$(\boldsymbol{x}, \boldsymbol{y}) \in K\text{-epi}\psi \quad \forall \boldsymbol{x} \in \psi^{-1}(T),\ \forall \boldsymbol{y} \in U(\boldsymbol{0}) + \overline{\boldsymbol{y}}.$$

由于 $\psi^{-1}(T)$ 和 $U(\boldsymbol{0}) + \overline{\boldsymbol{y}}$ 是非空开集，因此 $\mathrm{int}(K\text{-epi}\psi) \neq \varnothing$. ▯

引理 6.2.2 设 $S \subset \mathscr{X}$ 是开集，$\psi: S \to 2^{\mathscr{Y}}$ 是集值映射. 若对任意的 $\boldsymbol{x} \in S$，存在 $\boldsymbol{a} \in \mathscr{Y}$，使得 $\psi(\boldsymbol{x}) \subset \boldsymbol{a} - K$，则 $\mathrm{int}(K\text{-epi}\psi) \neq \varnothing$.

证明 设 $T = \boldsymbol{a} - K$，则 ψ 在 $\boldsymbol{x}$ 处是 K- 有界逆开的. 由引理 6.2.1，即得 $\mathrm{int}(K\text{-epi}\psi) \neq \varnothing$. ▯

定义 6.2.2 设集合 $S \subset \mathscr{X}$ 非空，点 $\boldsymbol{x}^0 \in S$，$\psi: S \to 2^{\mathscr{Y}}$ 是集值映射. 若存在映射 $\varphi: S \to \mathscr{Y}$，使得 $\varphi(\boldsymbol{x}) \in \psi(\boldsymbol{x})(\forall \boldsymbol{x} \in S)$，并且 φ 在点 $\boldsymbol{x}^0$ 的邻域 $U(\boldsymbol{x}^0)$ 内是连续的，则称 ψ 在点 $\boldsymbol{x}^0$ 处是连通的.

引理 6.2.3 设集合 $S \subset \mathscr{X}$ 非空，$\boldsymbol{x} \in S$，$\psi: S \to 2^{\mathscr{Y}}$ 是集值映射. 若 ψ 在 $\boldsymbol{x}$ 处是连通的，则 $\mathrm{int}(K\text{-epi}\psi) \neq \varnothing$.

证明 因为 φ 在点 $\boldsymbol{x}$ 的邻域 $U(\boldsymbol{x})$ 内是连续的，所以 $\varphi(U(\boldsymbol{x}))$ 是开集. 由此，我们有 $\big(U(\boldsymbol{x}),\ \varphi\big(U(\boldsymbol{x})\big)\big) \subset K\text{-epi}\psi$，故 $\mathrm{int}(K\text{-epi}\psi) \neq \varnothing$. ▯

以下给出锥凸集值映射的锥次微分的存在性定理.

定理 6.2.4 设 $S \subset \mathscr{X}$ 是内部非空的凸集，$\psi: S \to 2^{\mathscr{Y}}$ 是集值映射，$\psi(\boldsymbol{x}') \neq \varnothing \quad (\forall \boldsymbol{x}' \in S)$，并且 $\boldsymbol{x} \in \mathrm{int}S$，$\boldsymbol{p} \in \mathrm{int}K$. 又设 ψ 在 S 上是 K- 凸的，在 $\boldsymbol{x}$ 处是 K- 严格凸的和 K- 有界逆开的，并且 $\mathscr{E}\big(\psi(\boldsymbol{x}),\ K\big) \neq \varnothing$.

(1) 若 $\boldsymbol{y} \in \mathscr{E}\big(\psi(\boldsymbol{x}),\ K\big)$，则 $\partial\psi(\boldsymbol{x},\ \boldsymbol{y})_{\boldsymbol{p}} \neq \varnothing$.

(2) $\partial\psi(\boldsymbol{x})_{\boldsymbol{p}} \neq \varnothing$.

证明 (1) 首先证明 $K\text{-epi}\psi$ 是 $\mathscr{X} \times \mathscr{Y}$ 中的凸集. 设取任意的 $(\boldsymbol{x}^1,\ \boldsymbol{y}^1)$，$(\boldsymbol{x}^2,\ \boldsymbol{y}^2) \in K\text{-epi}\psi$，即 $\boldsymbol{y}^1 \in \psi(\boldsymbol{x}^1) + K$，$\boldsymbol{y}^2 \in \psi(\boldsymbol{x}^2) + K$. 因为 S 是凸的，设 $\lambda \in (0,\ 1)$，则 $\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2 \in S$. 又因为 K 也是凸的，ψ 在 S 上是 K- 凸的，依据定义 6.1.1 的(1) 有

$$\begin{aligned}\lambda\boldsymbol{y}^1 + (1-\lambda)\boldsymbol{y}^2 &\in \lambda\psi(\boldsymbol{x}^1) + (1-\lambda)\psi(\boldsymbol{x}^2) + K \\ &\subset \psi\big(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2\big) + K,\end{aligned}$$

从而 $\big(\lambda\boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2,\ \lambda\boldsymbol{y}^1 + (1-\lambda)\boldsymbol{y}^2\big) \in K\text{-epi}\psi$，故 $K\text{-epi}\psi$ 是凸集. 由 ψ 在 $\boldsymbol{x}$ 处是 K- 有界逆开的，根据引理 6.2.1 得知

$$\operatorname{int}(K\text{-epi}\psi) \neq \varnothing. \tag{6.2.2}$$

已知 $\boldsymbol{y} \in \mathscr{E}[\psi(\boldsymbol{x}), K]$，我们证明

$$(\boldsymbol{x}, \boldsymbol{y}) \notin \operatorname{int}(K\text{-epi}\psi). \tag{6.2.3}$$

事实上，若 $(\boldsymbol{x}, \boldsymbol{y}) \in \operatorname{int}(K\text{-epi}\psi)$，则存在点 $\boldsymbol{0}$ 的邻域 $U_1(\boldsymbol{0}) \subset \mathscr{Y}$ 使得

$$(\boldsymbol{x}, \boldsymbol{y}' + \boldsymbol{y}) \in K\text{-epi}\psi \quad \forall \boldsymbol{y}' \in U_1(\boldsymbol{0}). \tag{6.2.4}$$

因为 $\boldsymbol{p} \neq \boldsymbol{0}$ 和 $U_1(\boldsymbol{0})$ 是吸收集，所以存在 $\mu > 0$，使得 $-\mu\boldsymbol{p} \in U_1(\boldsymbol{0})$. 由此，从(6.2.4)得 $\boldsymbol{y}' - \mu\boldsymbol{p} \in \psi(\boldsymbol{x}) + K$，于是对应存在 $\overline{\boldsymbol{y}} \in \psi(\boldsymbol{x})$，$\overline{\boldsymbol{p}} \in K$ 使得 $\overline{\boldsymbol{y}} + \overline{\boldsymbol{p}} = \boldsymbol{y} - \mu\boldsymbol{p}$，即

$$\boldsymbol{y} - \overline{\boldsymbol{y}} = \overline{\boldsymbol{p}} + \mu\boldsymbol{p} \in \operatorname{int}K \subset K\backslash\{\boldsymbol{0}\}.$$

根据定义 6.1.2 的(1)，即知 $\boldsymbol{y} \notin \mathscr{E}(\psi(\boldsymbol{x}), K)$. 这导致与定理的条件相矛盾.

现在由(6.2.2)和(6.2.3)，根据定理 1.3.7 的(1)，可知存在 $(\boldsymbol{x}^*, \boldsymbol{y}^*) \in \mathscr{X}^* \times \mathscr{Y}^*$，$(\boldsymbol{x}^*, \boldsymbol{y}^*) \neq \boldsymbol{0}$，使得

$$\begin{gathered}\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle + \langle \boldsymbol{y}^*, \boldsymbol{y} \rangle \geqslant \langle \boldsymbol{x}^*, \boldsymbol{x}' \rangle + \langle \boldsymbol{y}^*, \boldsymbol{y}' \rangle \\ \forall \boldsymbol{x}' \in S, \forall \boldsymbol{y}' \in \psi(\boldsymbol{x}') + K.\end{gathered} \tag{6.2.5}$$

在(6.2.5)中 $\boldsymbol{y}^* \neq \boldsymbol{0}$，因为否则，若 $\boldsymbol{y}^* = \boldsymbol{0}$，由 $(\boldsymbol{x}^*, \boldsymbol{y}^*) \neq \boldsymbol{0}$ 知 $\boldsymbol{x}^* \neq \boldsymbol{0}$，代入(6.2.5)得到

$$\langle \boldsymbol{x}^*, \boldsymbol{x} - \boldsymbol{x}' \rangle \geqslant 0 \quad \forall \boldsymbol{x}' \in S. \tag{6.2.6}$$

于是，存在 $\overline{\boldsymbol{x}} \in \mathscr{X}$ 使得 $\mu = \langle \boldsymbol{x}^*, \overline{\boldsymbol{x}} \rangle > 0$. 再由 $\boldsymbol{x} \in \operatorname{int}S$ 知存在 $\boldsymbol{x}$ 的邻域 $U_2(\boldsymbol{x}) \subset \operatorname{int}S$，因此有充分大的 $M > 0$ 使 $\boldsymbol{x} + (\mu/M)\overline{\boldsymbol{x}} \in U_2(\boldsymbol{x})$. 将它代入(6.2.6)，有 $\langle \boldsymbol{x}^*, -(\mu/M)\overline{\boldsymbol{x}} \rangle \geqslant 0$，这导致 $-\mu^2 \geqslant 0$，产生矛盾. 现任取 $\boldsymbol{q} \in K$，因为 $\boldsymbol{y} \in \psi(\boldsymbol{x})$，故 $\boldsymbol{y} + \boldsymbol{q} \in \psi(\boldsymbol{x}) + K$，代入(6.2.5)有

$$\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle + \langle \boldsymbol{y}^*, \boldsymbol{y} \rangle \geqslant \langle \boldsymbol{x}^*, \boldsymbol{x} \rangle + \langle \boldsymbol{y}^*, \boldsymbol{y} + \boldsymbol{q} \rangle,$$

从而得

$$\langle -y^*, q\rangle \geqslant 0 \quad \forall q \in K.$$

注意到 $y^* \neq 0$，有 $-y^* \in K^*\setminus\{0\}$. 因为 $p \in \operatorname{int}K$，由定理 1.4.11 的(1) 得 $\langle -y^*, p\rangle > 0$. 令

$$\delta = \langle -y^*, p\rangle, \tag{6.2.7}$$

进一步证明下式成立：

$$\langle x^*, x\rangle + \langle y^*, y\rangle > \langle x^*, x'\rangle + \langle y^*, y'\rangle, \tag{6.2.8}$$

其中 $(x, y) \neq (x', y')(\forall x' \in S, \forall y' \in \psi(x') + K)$. 事实上，假若(6.2.8)不成立，从(6.2.5)可知存在$(\bar{x}, \bar{y}+\bar{q}) \in K\text{-epi}\psi$，$\bar{y} \in \psi(\bar{x})$ 和$(\bar{x}, \bar{y}+\bar{q}) \neq (x, y)$，使得

$$\langle x^*, x\rangle + \langle y^*, y\rangle = \langle x^*, \bar{x}\rangle + \langle y^*, \bar{y}\rangle + \bar{q}. \tag{6.2.9}$$

对任意的 $\lambda \in (0, 1)$ 令

$$x_\lambda = \lambda \bar{x} + (1-\lambda)x, \tag{6.2.10}$$

$$y_\lambda = \lambda \bar{y} + (1-\lambda)y, \tag{6.2.11}$$

因为 $\bar{y} \in \psi(\bar{x})$，$y \in \psi(x)$，$\psi$ 在 x 处是 K- 严格凸的，我们有

$$y_\lambda = \lambda \bar{y} + (1-\lambda)y \in \psi\big(\lambda \bar{x} + (1-\lambda)x\big) + \operatorname{int}K.$$

又因为 S 是凸集，故 $x_\lambda \in S$. 设 $y_\lambda = \eta + \gamma$，其中 $\eta \in \psi(x_\lambda)$，$\gamma \in \operatorname{int}K$. 代入(6.2.5)得

$$\langle x^*, x\rangle + \langle y^*, y\rangle \geqslant \langle x^*, x_\lambda\rangle + \langle y^*, \eta\rangle. \tag{6.2.12}$$

由(6.2.5)，(6.2.10)和(6.2.11)，推知有

$$\begin{aligned}
&\langle x^*, x_\lambda\rangle + \langle y^*, y_\lambda\rangle \\
&= \big\langle x^*, \lambda \bar{x} + (1-\lambda)x\big\rangle + \big\langle y^*, \lambda \bar{y} + (1-\lambda)y\big\rangle \\
&= \lambda\big[\langle x^*, \bar{x}\rangle + \langle y^*, \bar{y}\rangle\big]
\end{aligned}$$

$$+ (1-\lambda)\left[\langle x^*, x\rangle + \langle y^*, y\rangle\right]$$

$$= \langle x^*, x\rangle + \langle y^*, y\rangle - \lambda\langle y^*, \bar{q}\rangle.$$

据此，由(6.2.12)得知

$$\langle y^*, y_\lambda\rangle + \lambda\langle y^*, \bar{q}\rangle \geqslant \langle y^*, \eta\rangle.$$

再由 $\langle -y^*, \bar{q}\rangle \geqslant 0$，得到

$$\langle y^*, y_\lambda\rangle \geqslant \langle y^*, \eta\rangle. \qquad (6.2.13)$$

据定理 1.4.11 的(1)知 $\langle y^*, \gamma\rangle < 0$，从而有

$$\langle y^*, y_\lambda\rangle = \langle y^*, \eta + \gamma\rangle < \langle y^*, \eta\rangle,$$

它与(6.2.13)相矛盾.

最后，假设若 $\partial\psi(x, y)_p = \varnothing$，由定理 6.1.3 知，对于 $\delta > 0$ 有 $\partial\psi(x, y)_{\frac{p}{\delta}} = \varnothing$，即 $x^* \notin \partial\psi(x, y)_{\frac{p}{\delta}}$. 根据定义 6.1.3 的(1)，存在 $\bar{x} \in S$，$\bar{y} \in \psi(\bar{x})$，$(\bar{x}, \bar{y}) \neq (x, y)$，使得

$$y - \langle x^*, x\rangle \frac{p}{\delta} - (\bar{y} - \langle x^*, \bar{x}\rangle \frac{p}{\delta}) \in K\backslash\{0\}.$$

由 $-y^* \in K^*\backslash\{0\}$ 即可推得

$$\langle x^*, x\rangle + \langle y^*, y\rangle \leqslant \langle x^*, \bar{x}\rangle + \langle y^*, \bar{y}\rangle,$$

这与(6.2.6)相矛盾.

(2) 由已证明的(1)依据定义 6.1.3 的(1)，得知 $\partial\psi(x)_p \neq \varnothing$. ▯

设 $\mathscr{Y}$ 是向量空间.

定义 6.2.3 设集合 $S \subset \mathscr{Y}$ 非空，$K \subset \mathscr{Y}$ 是非平凡锥. 若 S 的每一个形如 $\{(y^\gamma - \mathrm{cl}K)^c \mid y^\gamma \in S\}_{\gamma\in\Gamma}$ 的开覆盖都有一个有限子覆盖(Γ 是指标集)，则称 S 是 K-半紧的.

引理 6.2.5 设集合 $Y \subset \mathscr{Y}$ 非空，$K \subset \mathscr{Y}$ 是尖闭凸锥. 若 Y 是 K-半紧的，则 $\mathscr{E}(Y, K) \neq \varnothing$.

证明 见文献[27]的定理 3.3.1.

定理 6.2.6 设 $S \subset \mathscr{X}$ 是非空开凸集，$\psi: S \to 2^{\mathscr{Y}}$ 是集值映射，$\psi(\boldsymbol{x}') \neq \varnothing$ $(\forall \boldsymbol{x}' \in S)$，$\boldsymbol{x} \in S$，$\boldsymbol{p} \in \mathrm{int}K$. 又设 ψ 在 S 上是 K-凸的，在 $\boldsymbol{x}$ 处是 K-严格凸的，并且存在 $\boldsymbol{a} \in \mathscr{Y}$ 使得对任意的 $\boldsymbol{x}' \in S$，有 $\psi(\boldsymbol{x}') \subset \boldsymbol{a} - K$，以及 $\psi(\boldsymbol{x})$是 K-半紧的.

(1) 若 $\boldsymbol{y} \in \mathscr{E}(\psi(\boldsymbol{x}), K)$，则 $\partial\, \psi(\boldsymbol{x}, \boldsymbol{y})_p \neq \varnothing$.

(2) $\partial\, \psi(\boldsymbol{x})_p \neq \varnothing$.

证明 由引理 6.2.2 和引理 6.2.5，以及定理 6.2.4 可证得. □

定理 6.2.7 设 $S \subset \mathscr{X}$ 是内部非空的凸集，$\psi: S \to 2^{\mathscr{Y}}$ 是集值映射，$\psi(\boldsymbol{x}') \neq \varnothing$ $(\forall \boldsymbol{x}' \in S)$，$\boldsymbol{x} \in \mathrm{int}S$，$\boldsymbol{p} \in \mathrm{int}K$. 又设 ψ 在 S 上是 K-凸的，在 $\boldsymbol{x}$ 处是 K-严格凸的和连通的，并且 $\psi(\boldsymbol{x})$是 K-半紧的.

(1) 若 $\boldsymbol{y} \in \mathscr{E}(\psi(\boldsymbol{x}), K)$，则 $\partial\psi(\boldsymbol{x}, \boldsymbol{y})_p \neq \varnothing$.

(2) $\partial\psi(\boldsymbol{x})_p \neq \varnothing$.

证明 由引理 6.2.3，引理 6.2.5 和定理 6.2.4，可以推得. □

以下给出锥凸集值映射的锥弱次微分的存在性定理.

定理 6.2.8 设 $S \subset \mathscr{X}$ 是内部非空的凸集，$\psi: S \to 2^{\mathscr{Y}}$ 是集值映射，$\psi(\boldsymbol{x}') \neq \varnothing$ $(\forall \boldsymbol{x}' \in S)$，$\boldsymbol{x} \in \mathrm{int}S$，$\boldsymbol{p} \in \mathrm{int}K$. 又设 ψ 在 S 上是 K-凸的，在 $\boldsymbol{x}$ 处是 K-有界逆开的，并且 $\mathscr{E}_w(\psi(\boldsymbol{x}), K) \neq \varnothing$.

(1) 若 $\boldsymbol{y} \in \mathscr{E}_w(\psi(\boldsymbol{x}), K)$，则 $\partial_w\psi(\boldsymbol{x}, \boldsymbol{y})_p \neq \varnothing$.

(2) $\partial_w\psi(\boldsymbol{x})_p \neq \varnothing$.

证明 (1) 类似于定理 6.2.4，可以证明 K-epiψ 是内部非空的凸集，并且对于 $\boldsymbol{y} \in \mathscr{E}_w(\psi(\boldsymbol{x}), K)$ 有 $(\boldsymbol{x}, \boldsymbol{y}) \notin \mathrm{int}(K\text{-epi}\psi)$. 据此，由定理 1.3.7 的(1) 可知，存在$(\boldsymbol{x}^*, \boldsymbol{y}^*) \in \mathscr{X}^* \times \mathscr{Y}^*$，$(\boldsymbol{x}^*, \boldsymbol{y}^*) \neq \boldsymbol{0}$，使得

$$\langle \boldsymbol{x}^*, \boldsymbol{x} \rangle + \langle \boldsymbol{y}^*, \boldsymbol{y} \rangle \geqslant \langle \boldsymbol{x}^*, \boldsymbol{x}' \rangle + \langle \boldsymbol{y}^*, \boldsymbol{y}' \rangle$$

$$\forall \boldsymbol{x}' \in S, \forall \boldsymbol{y}' \in \psi(\boldsymbol{x}') + K. \tag{6.2.14}$$

同定理 6.2.4 的证明，还可得 $\boldsymbol{y}^* \neq \boldsymbol{0}$，$-\boldsymbol{y}^* \in K^* \setminus \{\boldsymbol{0}\}$. 由于 $\boldsymbol{p} \in \mathrm{int}K$，根据定理 1.4.11 的(1) 有

$$\delta = \langle y^*, p\rangle > 0. \tag{6.2.15}$$

现在用反证法. 假设 $\partial_w \psi(x, y)_p = \varnothing$，则由定理 6.1.3 的(2)得知 $\partial_w \psi(x, y)_{\frac{p}{\delta}} = \varnothing$，于是由定义 6.1.3 的(2)可知存在 $\bar{x} \in S$，$\bar{y} \in \psi(\bar{x})$，使得

$$y - \langle x^*, x\rangle \frac{p}{\delta} - \left(\bar{y} - \langle x^*, \bar{x}\rangle \frac{p}{\delta}\right) \in \mathrm{int}K.$$

再根据定理 1.4.11，我们有

$$\left\langle -y^*, y - \bar{y} - \langle x^*, x - \bar{x}\rangle \frac{p}{\delta} \right\rangle > 0,$$

利用(6.2.15)整理上式，得

$$\langle x^*, x\rangle + \langle y^*, y\rangle < \langle x^*, \bar{x}\rangle + \langle y^*, \bar{y}\rangle,$$

这导致与(6.2.14)相矛盾.

(2) 由(1)依据定义 6.1.3 的(2)，即得 $\partial_w \psi(x)_p \neq \varnothing$. ▯

定理 6.2.9　设 $S \subset \mathscr{X}$ 是非空开凸集，$\psi: S \to 2^{\mathscr{Y}}$ 是集值映射，$\psi(x') \neq \varnothing$ $(\forall x' \in S)$，$x \in S$，$p \in \mathrm{int}K$. 又设 ψ 在 S 上是 K-凸的，存在 $a \in \mathscr{Y}$ 使得对任意的 $x' \in S$ 有 $\psi(x') \subset a - K$，并且 $\psi(x')$ 是 K 半紧的.

(1) 若 $y \in \mathscr{E}_w(\psi(x), K)$，则 $\partial_w \psi(x, y)_p \neq \varnothing$.

(2) $\partial_w \psi(x)_p \neq \varnothing$.

证明　由引理 6.2.2 可知 $\mathrm{int}(K\text{-epi}\psi) \neq \varnothing$，又由引理 6.2.5 有 $\mathscr{E}_w(\psi(x), K) \neq \varnothing$. 于是，由定理 6.2.8 即得结论. ▯

定理 6.2.10　设 $S \subset \mathscr{X}$ 是内部非空的凸集，$\psi: S \to 2^{\mathscr{Y}}$ 是集值映射，$\psi(x') \neq \varnothing$ $(\forall x' \in S)$，$x \in \mathrm{int}S$，$p \in \mathrm{int}K$. 又设 ψ 在 S 上是 K-凸的，在 x 处是连通的，并且 $\psi(x)$是 K-半紧的.

(1) 若 $y \in \mathscr{E}_w(\psi(x), K)$，则 $\partial_w \psi(x, y)_p \neq \varnothing$.

(2) $\partial_w \psi(x)_p \neq \varnothing$.

证明　由引理 6.2.3、引理 6.2.5 和定理 6.2.8 可得证. ▯

下面给出锥次微分和锥弱次微分存在的充分条件.

定理6.2.11 设集合$S\subset\mathscr{X}$非空，$x\in S$，$\psi: S\to 2^{\mathscr{Y}}$是集值映射，$y\in\psi(x)$. 又$K^*\subset\mathscr{Y}^*$是$K$的对偶锥，$x^*\in\mathscr{X}^*$，$p\in\mathrm{int}K$.

(1) 若存在$q^*\in\mathrm{int}K^*$使得

$$\langle q^*, y'-y\rangle\geqslant\langle x^*, x'-x\rangle\langle q^*, p\rangle$$
$$\forall x'\in S, \forall y'\in\psi(x'), \tag{6.2.16}$$

则$x^*\in\partial\psi(x, y)_p$.

(2) 若存在$q^*\in K^*\backslash\{0\}$使得

$$\langle q^*, y'-y\rangle\geqslant\langle x^*, x'-x\rangle\langle q^*, p\rangle$$
$$\forall x'\in S, \forall y'\in\psi(x'), \tag{6.2.17}$$

则$x^*\in\partial_w\psi(x, y)_p$.

证明 (1) 反之，假设$x^*\notin\partial\psi(x, y)_p$，由定义6.1.3的(1)知存在$\bar{x}\in S$，$\bar{y}\in\psi(\bar{x})$使得

$$y-\langle x^*, x\rangle p-(\bar{y}-\langle x^*, \bar{x}\rangle p)\in K\backslash\{0\}.$$

根据定理1.4.11的(2)有

$$\Big\langle q^*, y-\langle x^*, x\rangle p-(\bar{y}-\langle x^*, \bar{x}\rangle p)\Big\rangle>0,$$

整理上式得

$$\langle q^*, y-\bar{y}\rangle>\langle x^*, x-\bar{x}\rangle\langle q^*, p\rangle,$$

这导致与(6.2.16)矛盾.

(2) 用反证法. 利用定理1.4.11，类似于(1)的证明可得证. □

下面的定理说明，在凸性条件下，(6.2.17)还是锥次微分和锥弱次微分存在的必要条件.

定理6.2.12 设集合$S\subset\mathscr{X}$非空，$x\in S$，$\psi: S\to 2^{\mathscr{Y}}$是集值映射，$y\in\psi(x)$，$x^*\in\mathscr{X}^*$，$p\in\mathrm{int}K$，$K^*$是$K$的对偶锥. 若$\psi$在$S$上是$K$-凸的，并且$x^*\in\partial_w\psi(x, y)_p$，则存在$q^*\in K^*\backslash\{0\}$，使得

$$\langle q^*, y'-y\rangle \geqslant \langle x^*, x'-x\rangle\langle q^*, p\rangle$$

$$\forall x' \in S,\ \forall y' \in \psi(x').$$

证明　作集合

$$E = \{y' \in \mathscr{Y} \mid y' = y - \bar{y} + \langle x^*, \bar{x} - x\rangle p - \bar{q},$$

$$\bar{x} \in S,\ \bar{y} \in \psi(\bar{x}),\ \bar{q} \in K\}, \tag{6.2.18}$$

则由 S 和 K 是凸的,以及 ψ 是 K-凸的可以推知 E 是凸集. 下面证明

$$\operatorname{int}K \cap E = \varnothing. \tag{6.2.19}$$

事实上,若存在 $q' \in \operatorname{int}K$ 使 $q' \in E$,由(6.2.18)则有

$$q' = y - \bar{y} + \langle x^*, \bar{x} - x\rangle p - \bar{q},$$

即

$$y - \langle x^*, x\rangle p - (\bar{y} - \langle x^*, \bar{x}\rangle p) = \bar{q} + q' \in \operatorname{int}K.$$

于是,按定义 6.1.3 的(2)得知 $x^* \notin \partial_w\psi(x, y)_p$,它与所设矛盾. 现在由(6.2.19)根据定理 1.3.7 可知,存在 $q^* \in \mathscr{Y}^*\backslash\{\theta\}$,使得

$$\langle q^*, q\rangle \geqslant \langle q^*, y'\rangle \quad \forall q \in K,\ \forall y' \in E.$$

由 $\theta \in K \cap E$,从上式得知对任意的 $q \in K$ 有 $\langle q^*, q\rangle \geqslant 0$,故 $q^* \in K^*\backslash\{\theta\}$. 再由 $0 \geqslant \langle q^*, y'\rangle$ $(\forall y' \in E)$,即得(6.2.17). ▯

特别地,对任意的 $q^* \in K^*\backslash\{\theta\}$,作集值函数 $\theta: S \to 2^R$, $x \mapsto \theta(x)$, $\theta(x) = \langle q^*, \psi(x)\rangle = \{\langle q^*, y\rangle \mid y \in \psi(x)\}$ $(\forall x \in S)$. 取 R 中的尖闭凸锥为 R_+,显然因 $R_+\backslash\{\theta\} = \operatorname{int}R_+$,再令 $p = 1$,则由定义 6.1.3 可知,对于 $x \in S$, $y \in \psi(x)$,我们有

$$\partial\theta(x, \langle q^*, y\rangle)_1 = \partial_w\theta(x, \langle q^*, y\rangle)_1.$$

定理 6.2.13　设集合 $S \subset \mathscr{X}$ 非空,$x \in S$, $\psi: S \to 2^{\mathscr{Y}}$ 是集值映射,$y \in \psi(x)$. 又设 K^* 是 K 的对偶锥,ψ 在 S 上是 K-凸的,$p \in \operatorname{int}K$.

(1) $\bigcup\limits_{q^* \in \mathrm{int}K^*} \frac{1}{\langle q^*, p\rangle}\partial\theta\big(x, \langle q^*, y\rangle\big)_1 \subset \partial\psi(x, y)_p.$

(2) $\bigcup\limits_{q^* \in K^* \setminus \{\theta\}} \frac{1}{\langle q^*, p\rangle}\partial\theta\big(x, \langle q^*, y\rangle\big)_1 = \partial_w\psi(x, y)_p.$

证明 (1) 对任意的 $q^* \in \mathrm{int}K^*$,设

$$x^* \in \big(\langle q^*, p\rangle\big)^{-1}\partial\theta\big(\bar{x}, \langle q^*, y\rangle\big)_1,$$

则

$$\langle q^*, p\rangle x^* \in \partial\theta\big(x, \langle q^*, y\rangle\big)_1.$$

由 ψ 在 S 上是 K-凸的,容易证明 θ 在 S 上是 R_+-凸的. 于是,由定理 6.2.12 知存在 $\lambda \in R_+ \setminus \{\boldsymbol{0}\}$,使得

$$\lambda\big(\langle q^*, y' - \langle q^*, y\rangle\big) \leqslant \big\langle\langle q^*, p\rangle x^*, x' - x\big\rangle\lambda$$

$$\forall x' \in S, \forall \in \theta(x'),$$

即

$$\langle q^*, y' - y\rangle \geqslant \langle x^*, x' - x\rangle\langle q^*, p\rangle$$

$$\forall x' \in S, \forall y' \in \psi(x').$$

由此,根据定理 6.2.11 得到 $x^* \in \partial\psi(x, y)_p$.

(2) 类似于(1)中的证明,由定理 6.2.11 和定理 6.2.12 可以推得

$$\bigcup_{q^* \in K^* \setminus \{\theta\}} \frac{1}{\langle q^*, p\rangle}\partial\theta\big(x, \langle q^*, y\rangle\big)_1 \subset \partial_w\psi(x, y)_p.$$

其次,设 $x^* \in \partial_w\psi(x, y)_p$,由定理 6.2.12 知存在 $q^* \in K^* \setminus \{\boldsymbol{0}\}$,使得

$$\langle q^*, y' - y\rangle \geqslant \langle x^*, x' - x\rangle\langle q^*, p\rangle$$

$$\forall x' \in S, \forall y' \in \psi(x').$$

于是,根据定理 6.2.11 得 $x^*\langle q^*, p\rangle \in \partial\theta\big(x, \langle q^*, y\rangle\big)_1$,即有 $x^* \in \big(\langle q^*, p\rangle\big)^{-1}\partial\theta\big(x, \langle q^*, y\rangle\big)_1$. □

设 $f: S \to R$ 是实值函数，$f(x) = \inf\langle \boldsymbol{q}^*, \psi(\boldsymbol{x})\rangle$，其中 $\boldsymbol{q}^* \in K^* \setminus \{\boldsymbol{0}\}$.

定理 6.2.14　设 $S \subset \mathscr{X}$ 是非空凸集，$\psi: S \to 2^{\mathscr{Y}}$ 是集值映射，$\boldsymbol{x} \in S$，$\boldsymbol{y} \in \psi(\boldsymbol{x})$，并且 $\boldsymbol{p} \in \operatorname{int}K$，$K^*$ 是 K 的对偶锥. 若 ψ 在 S 上是 K-凸的，$f(\boldsymbol{x}) = \langle \boldsymbol{q}^*, \boldsymbol{y}\rangle\ (\forall \boldsymbol{q}^* \in K^* \setminus \{\boldsymbol{0}\})$，对任意的 $\boldsymbol{x}' \in S$，$\psi(\boldsymbol{x}')$ 是非空紧集，则对任意的 $\boldsymbol{q}^* \in K^* \setminus \{\boldsymbol{0}\}$，函数 f 在 S 上是凸的，并且

(1) $\displaystyle\bigcup_{\boldsymbol{q}^* \in \operatorname{int}K^*} \frac{1}{\langle \boldsymbol{q}^*, \boldsymbol{p}\rangle} \partial f(\boldsymbol{x}) \subset \partial\, \psi(\boldsymbol{x}, \boldsymbol{y})_p$,

(2) $\displaystyle\bigcup_{\boldsymbol{q}^* \in K^* \setminus \{\boldsymbol{0}\}} \frac{1}{\langle \boldsymbol{q}^*, \boldsymbol{p}\rangle} \partial f(\boldsymbol{x}) = \partial_w \psi(\boldsymbol{x}, \boldsymbol{y})_p$,

其中 $\partial f(\boldsymbol{x}) = \partial f(\boldsymbol{x}, \langle \boldsymbol{q}^*, \boldsymbol{y}\rangle)_1$.

证明　任取 $\boldsymbol{x}^1, \boldsymbol{x}^2 \in S$，$\lambda \in (0, 1)$，因为已知对任何连续线性泛函 $\boldsymbol{q}^* \in K^* \setminus \{\boldsymbol{0}\}$，$\psi(\boldsymbol{x})\ (\forall \boldsymbol{x} \in S)$ 是紧集，所以存在 $\boldsymbol{y}^1 \in \psi(\boldsymbol{x}^1)$，$\boldsymbol{y}^2 \in \psi(\boldsymbol{x}^2)$，使得 $f(\boldsymbol{x}^1) = \langle \boldsymbol{q}^*, \boldsymbol{y}^1\rangle$，$f(\boldsymbol{x}^2) = \langle \boldsymbol{q}^*, \boldsymbol{y}^2\rangle$. 由 S 是凸的，ψ 在 S 上是 K-凸的，则有

$$\lambda \boldsymbol{y}^1 + (1 - \lambda)\boldsymbol{y}^2 \in \psi(\lambda \boldsymbol{x}^1 + (1 - \lambda)\boldsymbol{x}^2) + K.$$

因此，存在 $\boldsymbol{y}_\lambda \in \psi(\lambda \boldsymbol{x}^1 + (1 - \lambda)\boldsymbol{x}^2)$ 和 $\boldsymbol{q} \in K$，使得

$$\lambda \boldsymbol{y}^1 + (1 - \lambda)\boldsymbol{y}^2 = \boldsymbol{y}_\lambda + \boldsymbol{q}.$$

据此，由 $\boldsymbol{q}^* \in K^* \setminus \{\boldsymbol{0}\}$ 得知

$$\langle \boldsymbol{q}^*, \lambda \boldsymbol{y}^1\rangle + \langle \boldsymbol{q}^*, (1 - \lambda)\boldsymbol{y}^2\rangle = \langle \boldsymbol{q}^*, \boldsymbol{y}_\lambda\rangle + \langle \boldsymbol{q}^*, \boldsymbol{q}\rangle,$$

即

$$\begin{aligned}\lambda f(\boldsymbol{x}^1) + (1 - \lambda) f(\boldsymbol{x}^2) &\geqslant \langle \boldsymbol{q}^*, \boldsymbol{y}_\lambda\rangle \\ &\geqslant f(\lambda \boldsymbol{x}^1 + (1 - \lambda)\boldsymbol{x}^2),\end{aligned}$$

故 f 在 S 上是凸的. 由定理 6.2.11 和定理 6.2.12 有

$$\partial f(\boldsymbol{x}) = \{\boldsymbol{x}^* \mid \langle \boldsymbol{x}^*, \boldsymbol{x}' - \boldsymbol{x}\rangle \leqslant f(\boldsymbol{x}') - f(\boldsymbol{x})\ \forall \boldsymbol{x}' \in S\}. \tag{6.2.20}$$

现在证明 $\partial f(x)=\partial\theta\left(x,\langle q^*,y\rangle\right)_1$. 设 $x^*\in\partial f(x)$,由(6.2.20)和 f 的定义可知,对任意的 $x'\in S$ 和任意的 $y'\in\psi(x')$ 有

$$\langle q^*,y'\rangle-\langle q^*,y\rangle\geqslant f(x')-f(x)\geqslant\langle x^*,x'\rangle.$$

由此,根据定理 6.2.11 得 $x^*\in\partial\theta\left(x,\langle q^*,y\rangle\right)_1$. 反之,设 $x^*\in\partial\theta\left(x,\langle q^*,y\rangle\right)_1$,由定理6.2.12知对任意的$x'\in S$和任意的$y'\in\psi(x')$ 有

$$\langle q^*,y'\rangle-\langle q^*,y\rangle\geqslant\langle x^*,x'-x\rangle.$$

注意到 $\psi(x')$是紧集,则存在 y_x 使得 $f(x')=\langle q^*,y_x\rangle$, $y_x\in\psi(x')$,故由上式得知

$$\langle x^*,x'-x\rangle\leqslant f(x')-f(x)\quad\forall x'\in S,$$

再由(6.2.20)便得 $x^*\in\partial f(x)$. 根据以上结果,由定理 6.2.13,即得(1)和(2)的结论.□

对于 Banach 空间的情况,有以下一些结果.

定理 6.2.15 设 $\mathscr{B}$ 是 Banach 空间,$S\subset\mathscr{X}$ 是非空凸集,ψ: $S\to 2^{\mathscr{B}}$ 是集值映射,$\psi(x')\neq\varnothing$ $(\forall x'\in S)$. 又设 $x\in S$, $y\in\psi(x)$, $p\in\operatorname{int}K$, K^* 是 K 的对偶锥. 若 ψ 在 S 上是 K-凸的,则

(1) $\displaystyle\bigcup_{q^*\in K_1^*}\frac{1}{\langle q^*,p\rangle}\partial\theta\left(x,\langle q^*,y\rangle\right)_1\subset\partial\psi(x,y)_p$,

(2) $\displaystyle\bigcup_{q^*\in K_2^*}\frac{1}{\langle q^*,p\rangle}\partial\theta\left(x,\langle q^*,y\rangle\right)_1=\partial_w\psi(x,y)_p$,

其中
$$K_1^*=\left\{q^*\in\operatorname{int}K^*\mid\|q^*\|_*=1\right\},$$
$$K_2^*=\left\{q^*\in K^*\setminus\{0\}\mid\|q^*\|_*=1\right\},$$

$\|\cdot\|_*$ 是对偶空间 $\mathscr{B}^*$ 的范数.

证明 (1) 由定理 6.2.13 结论显然成立.

(2) 据定理 6.2.13 的(2),只要证明

$$\partial_w\psi(x,y)_p\subset\bigcup_{q^*\in K_2^*}\frac{1}{\langle q^*,p\rangle}\partial\theta\left(x,\langle q^*,y\rangle\right)_1.$$

设 $x^* \in \partial_w \psi(x, y)_p$，由定理 6.2.12，存在 $q^* \in K^* \setminus \{0\}$ 使得

$$\langle q^*, y' - y\rangle \geqslant \langle x^*, x' - x\rangle \langle q^*, p\rangle$$

$$\forall x' \in S, \forall y' \in \psi(x').$$

因为 $\|q^*\|_* \neq 0$，在上式两端同除一个 $\|q^*\|_*$，令 $q_1^* = q^* / \|q^*\|_*$，得

$$\langle q_1^*, y' - y\rangle \geqslant \langle x^*, x' - x\rangle \langle q_1^*, p\rangle$$

$$\forall x' \in S, \forall y' \in \psi(x').$$

因此，得知 $q_1^* \in K_2^*$，$x^* \langle q_1^*, p\rangle \in \partial\, \theta(x, \langle q_1^*, y\rangle)_1$. 定理得证. □

定理 6.2.16　设 $\mathscr{B}$ 是 Banach 空间，$S \subset \mathscr{X}$ 是非空凸集，ψ: $S \to 2^{\mathscr{B}}$ 是集值映射. 又设 $x \in S$，$y \in \psi(x)$，$p \in \text{int}K$，K^* 是 K 的对偶锥. 若 ψ 在 S 上是 K-凸的，$f(x) = \langle q^*, y\rangle (\forall q^* \in K^* \setminus \{0\})$，对任意的 $x' \in S$，$\psi(x')$ 是紧集，则

(1) $\displaystyle\bigcup_{q^* \in K_1^*} \frac{1}{\langle q^*, p\rangle} \partial f(x) \subset \partial \psi(x, y)_p$，

(2) $\displaystyle\bigcup_{q^* \in K_2^*} \frac{1}{\langle q^*, p\rangle} \partial f(x) = \partial_w \psi(x, y)_p$，

其中　$\partial f(x) = \partial f(x, \langle q^*, y\rangle)_1$，

$$K_1^* = \{q^* \in \text{int}K^* \mid \|q^*\|_* = 1\},$$

$$K_2^* = \{q^* \in K^* \setminus \{0\} \mid \|q^*\|_* = 1\}.$$

证明　由定理 6.2.14 和定理 6.2.15 可得证. □

我们可以减弱定理 6.2.14 和定理 6.2.16 的条件，分别给出如下结果.

定理 6.2.17　设 $S \subset \mathscr{X}$ 是非空凸集，ψ: $S \to 2^{\mathscr{Y}}$ 是集值映射，$x \in S$，$y \in \psi(x)$，$p \in \text{int}K$，K^* 是 K 的对偶锥. 若 ψ 在 S 上是 K-凸的，对任意的 $x' \in S$，$\psi(x)$ 是非空紧集，则 f 在 S 上是

凸的.

(1) 再若存在 $q^* \in \mathrm{int}K^*$ 使得 $f(x)=\langle q^*, y\rangle$, 则

$$\{\langle q^*, p\rangle\}^{-1}\partial f[x, \langle q^*, y\rangle]_1 \subset \partial\psi(x, y)_p.$$

(2) 再若存在 $q^* \in K^*\setminus\{0\}$ 使得 $f(x)=\langle q^*, y\rangle$,则

$$\{\langle q^*, p\rangle\}^{-1}\partial f[x, \langle q^*, y\rangle]_1 \subset \partial_w\psi(x, y)_p.$$

(3) $\partial_w(x, y)_p \subset \bigcup\limits_{q^*\in K^*\setminus\{0\}} \dfrac{1}{\langle q^*, p\rangle}\partial f[x, \langle q^*, y\rangle]_1$.

证明 类似于定理 6.2.14 的证明可知,f 在 S 上是凸的,并且 $\partial f(x)=\partial\theta[x, \langle q^*, y\rangle]_1$.

(1) 当 $q^*\in \mathrm{int}K^*$, $x^*\in\{\langle q^*, p\rangle\}^{-1}\partial\theta[x, \langle q^*, y\rangle]_1$,类似于定理 6.2.13 的(1) 的证明,可以推知 $x^*\in\partial\psi(x, y)_p$.

(2) 与(1)的证明类似可得到.

(3) 由定理 6.2.12 知当 $x^*\in\partial_w\psi(x, y)_p$ 时,则存在 $q^*\in K^*\setminus\{0\}$使得

$$\langle q^*, y'-y\rangle \geqslant \langle x^*, x'-x\rangle\langle q^*, p\rangle$$

$$\forall x'\in S, \forall y'\in\psi(x'). \qquad (6.2.21)$$

设集值函数 $\theta(x)=\langle q^*, \psi(x)\rangle$,在(6.2.21) 中令 $x'=x$ 得

$$\langle q^*, y'\rangle \geqslant \langle q^*, y\rangle, \quad \forall y'\in\psi(x'),$$

故有 $f(x)=\langle q^*, y\rangle$. 再由(6.2.21) 得到

$$f(x')-f(x) \geqslant \langle\langle q^*, p\rangle x^*, x'-x\rangle \quad \forall x'\in S,$$

所以 $\langle q^*, p\rangle x^*\in\partial f[x, \langle q^*, y\rangle]_1$,这就证明了(3). ▯

定理 6.2.18 设 $\mathscr{Y}$ 是 Banach 空间, $S\subset\mathscr{X}$ 是非空凸集,ψ: $S\to 2^{\mathscr{Y}}$ 是集值映射, $x\in S$, $y\in\psi(x)$, $p\in\mathrm{int}K$, K^* 是 K 的对偶锥. 若 ψ 在 S 上是 K-凸的,对任意的 $x'\in S$, $\psi(x')$ 是非空紧集,则

$$\partial_w \psi(x, y) \subset \bigcup_{q^* \in K_2^*} \frac{1}{\langle q^*, p\rangle} \partial f\big(x, \langle q^*, y\rangle\big)_p,$$

其中 $K_2^* = \{q^* \in K^* \setminus \{\boldsymbol{0}\} \mid \| q^* \|_* = 1\}$.

证明　与定理 6.2.17 中(3)的证明类似. □

定理 6.2.17 提供了我们求锥次梯度和锥弱次梯度的途径.

例 6.2.1　设集合 $S \subset \mathscr{X}$ 非空,对任意的 $x \in S$, $\psi(x) = T \subset \mathscr{Y}$ 是非空紧集. 考虑集值映射 $\psi: S \to 2^{\mathscr{Y}}$, $x \mapsto \psi(x)$. 这时,对任何 $q^* \in K^* \setminus \{\boldsymbol{0}\}$, $f(x) = \inf_{y \in T} \langle q^*, y\rangle$ 是一个常数,所以 $\partial f(x) = \{\boldsymbol{0}\}$. 由定理 6.2.17 和定理 6.2.18,得 $\partial \psi(x)_p = \{\boldsymbol{0}\}$ ($\forall x \in S$).

例 6.2.2　设集合 $S \subset \mathscr{X}$ 非空,$\psi = f: S \to R^m$ 是实向量函数,取 $K = R_+^m$,并且 f 是 R_+^m - 凸的可微函数. 记 $f(x) = \big(f_1(x), \cdots, f_m(x)\big)^T \in R^m (x \in S)$,令 $p = (1, \cdots, 1)^T \in R^m$, $q^* = (q_1, \cdots, q_m)^T \in R_+^m \setminus \{\boldsymbol{0}\}$. $\langle q^*, p\rangle = \sum_{i=1}^m q_i$. 设 $x \in S$, $y = f(x)$, $g(x) = \langle q^*, y\rangle$,则有

$$\partial g\big(x, \langle q^*, y\rangle\big)_1 = \sum_{i=1}^m q_i \nabla f_i(x).$$

由定理 6.2.17 得到

$$\partial f(x, y)_p \supset \bigcup_{q^* \in \operatorname{int} K^*} \frac{1}{\sum_{i=1}^m q_i} \sum_{i=1}^m q_i \nabla f_i(x)$$

和

$$\partial_w f(x, y)_p = \bigcup_{q^* \in K^* \setminus \{\boldsymbol{0}\}} \frac{1}{\sum_{i=1}^m q_i} \sum_{i=1}^m q_i \nabla f_i(x),$$

即

$$\partial f(x)_p \supset \left\{ \sum_{i=1}^m q_i \nabla f_i(x) \,\middle|\, \sum_{i=1}^m q_i = 1, q_i \in (0, 1) \right\}$$

和

$$\partial_w f(x)_p = \left\{ \sum_{i=1}^{m} q_i \nabla f_i(x) \,\middle|\, \sum_{i=1}^{m} q_i = 1,\ q_i \in [0,1] \right\}.$$

由此可知道，f 的分量梯度的凸组合构成了 f 的 K-次微分.

例 6.2.3 考虑实值函数 $g: S \times Y \to R$，其中 S 和 Y 分别是 Banach 空间 $\mathscr{B}$ 和 $\mathscr{D}$ 中的非空集合. 作集值函数 $\psi: S \to 2^R$，$x \mapsto \psi(x) = \{t \mid t = g(x, y),\ y \in Y\}$，设 $f(x) = \inf_{y \in Y} g(x, y)$ $(\forall x \in S)$. 令 $p = 1$，$q^* = 1$，且有 $f(x) = \inf \psi(x)$. 如果 g 在 $S \times Y$ 上是下半连续的正常凸函数，并且当 $x \in S$，$y \in Y$ 时有 $f(x) = g(x, y)$，则

$$x^* \in \partial f(x) \Longleftrightarrow (x^*, 0) \in \partial \psi(x, y)_1.$$

设 $t = g(x, y)$，由定理 6.2.17 得到

$$\partial \psi(x, t)_1 = \partial_w \psi(x, t)_1 = \partial f(x).$$

例 6.2.4 考虑实值向量函数 $f: S \times Y \to R^m$，其中 S 和 Y 分别是 Banach 空间 $\mathscr{B}$ 和 $\mathscr{D}$ 中的非空集合. 设 $p = (1, \cdots, 1)^T \in R^m$，$q^* \in R_+^m \setminus \{0\}$，函数 $g: S \times Y \to R$，$g(x, y) = \langle q^*, f(x, y) \rangle$ $(\forall (x, y) \in S \times Y)$. 作集值函数 $\psi: S \to 2^{R^m}$，$x \mapsto \psi(x) = \{r \mid r = f(x, y),\ y \in Y\}$. 令

$$\bar{f}(x) = \inf_{y \in Y} g(x, y) = \inf_{y \in Y} \langle q^*, f(x, y) \rangle,$$

设 f 在 $S \times Y$ 上是下半连续的正常凸函数，由例 6.2.3 和定理 6.2.17 可得

$$x^* \in \partial \bar{f}(x) \Longleftrightarrow (x^*, 0) \in \partial g(x, y) = \sum_{i=1}^{m} q_i^* \partial f_i(x, y),$$

其中 $q^* = (q_1^*, \cdots, q_m^*)^T \in R_+^m \setminus \{0\}$，$f = (f_1, \cdots, f_m)^T \in R^m$. 此外，有

$$\partial\psi(\boldsymbol{x},\boldsymbol{r})_{\boldsymbol{p}} \supset \left\{\boldsymbol{x}^* \in \mathscr{X}^* \,\middle|\, (\boldsymbol{x}^*,\boldsymbol{0}) \in \sum_{i=1}^{m} q_i^* \partial f_i(\boldsymbol{x},\boldsymbol{y}),\right.$$

$$\left.\sum_{i=1}^{m} q_i^* = 1,\ q_i^* \in (0,1),\ i=1,\cdots,m\right\},$$

$$\partial_w\psi(\boldsymbol{x},\boldsymbol{r})_{\boldsymbol{p}} = \left\{\boldsymbol{x}^* \in \mathscr{X}^* \,\middle|\, (\boldsymbol{x}^*,\boldsymbol{0}) \in \sum_{i=1}^{m} q_i^* \partial f_i(\boldsymbol{x},\boldsymbol{y}),\right.$$

$$\left.\sum_{i=1}^{m} q_i^* = 1,\ q_i^* \in [0,1],\ i=1,\cdots,m\right\}.$$

注意，上式成立的必要条件是：存在 $\boldsymbol{q}^*$ 使得 $\overline{f}(\boldsymbol{x}) = \langle \boldsymbol{q}^*, \boldsymbol{r}\rangle = g(\boldsymbol{x},\boldsymbol{y})$.

§6.3　基本特性

现在讨论锥弱次微分及其相应映射的若干基本特性，如凸性、闭性、半连续性、连通性和次单调性等.

若无特别说明，本节均设 $S \subset \mathscr{X}$ 是非空凸集，$\psi: S \to 2^{\mathscr{Y}}$，$\boldsymbol{x} \mapsto \psi(\boldsymbol{x})$ 是集值映射，对任意的 $\boldsymbol{x} \in S$ 而言，集合 $\psi(\boldsymbol{x}) \subset \mathscr{Y}$ 非空，$K \subset \mathscr{Y}$ 是内部非空的尖闭凸锥.

先介绍锥弱次微分的凸性和闭性.

定理 6.3.1　设 $\boldsymbol{x} \in S$，$\boldsymbol{y} \in \psi(\boldsymbol{x})$，$\boldsymbol{p} \in \mathrm{int}K$. 若 ψ 在 S 上是 K-凸的，则 $\partial_w\psi(\boldsymbol{x},\boldsymbol{y})_{\boldsymbol{p}}$ 是凸集.

证明　设 $\boldsymbol{x}_1^*, \boldsymbol{x}_2^* \in \partial_w\psi(\boldsymbol{x},\boldsymbol{y})_{\boldsymbol{p}}$，对任意的 $\lambda \in (0,1)$，记 $\boldsymbol{x}_\lambda^* = \lambda\boldsymbol{x}_1^* + (1-\lambda)\boldsymbol{x}_2^*$. 为证 $\boldsymbol{x}_\lambda^* \in \partial_w\psi(\boldsymbol{x},\boldsymbol{y})_{\boldsymbol{p}}$，用反证法，假设 $\boldsymbol{x}_\lambda^* \notin \partial_w\psi(\boldsymbol{x},\boldsymbol{y})_{\boldsymbol{p}}$，则按定义 6.1.3 可知，存在 $\overline{\boldsymbol{x}} \in S$，$\overline{\boldsymbol{y}} \in \psi(\overline{\boldsymbol{x}})$，$(\overline{\boldsymbol{x}},\overline{\boldsymbol{y}}) \neq (\boldsymbol{x},\boldsymbol{y})$，有

$$\boldsymbol{y} - \langle \boldsymbol{x}_\lambda^*, \boldsymbol{x}\rangle \boldsymbol{p} - \left(\overline{\boldsymbol{y}} - \langle \boldsymbol{x}_\lambda^*, \overline{\boldsymbol{x}}\rangle \boldsymbol{p}\right) \in \mathrm{int}K. \tag{6.3.1}$$

由 $\boldsymbol{x}_1^*, \boldsymbol{x}_2^* \in \partial_w\psi(\boldsymbol{x},\boldsymbol{y})_{\boldsymbol{p}}$，根据定理 6.2.12 知，存在 $\boldsymbol{q}_1^*, \boldsymbol{q}_2^* \in K^*\setminus\{\boldsymbol{0}\}$ 使

$$\langle \boldsymbol{q}_1^*, \bar{\boldsymbol{y}} - \boldsymbol{y}\rangle \geqslant \langle \boldsymbol{x}_1^*, \bar{\boldsymbol{x}} - \boldsymbol{x}\rangle\langle \boldsymbol{q}_1^*, \boldsymbol{p}\rangle, \qquad (6.3.2)$$

$$\langle \boldsymbol{q}_2^*, \bar{\boldsymbol{y}} - \boldsymbol{y}\rangle \geqslant \langle \boldsymbol{x}_2^*, \bar{\boldsymbol{x}} - \boldsymbol{x}\rangle\langle \boldsymbol{q}_2^*, \boldsymbol{p}\rangle. \qquad (6.3.3)$$

从(6.3.1)有

$$\langle \boldsymbol{q}_1^*, \boldsymbol{y} - \bar{\boldsymbol{y}}\rangle + \langle \boldsymbol{x}_\lambda^*, \bar{\boldsymbol{x}} - \boldsymbol{x}\rangle \boldsymbol{p} > 0, \qquad (6.3.4)$$

$$\langle \boldsymbol{q}_2^*, \boldsymbol{y} - \bar{\boldsymbol{y}}\rangle + \langle \boldsymbol{x}_\lambda^*, \bar{\boldsymbol{x}} - \boldsymbol{x}\rangle \boldsymbol{p} > 0. \qquad (6.3.5)$$

由 (6.3.2) + (6.3.4) 和(6.3.3) + (6.3.5),分别可得

$$\langle \boldsymbol{x}_\lambda^*, \bar{\boldsymbol{x}} - \boldsymbol{x}\rangle > \langle \boldsymbol{x}_1^*, \bar{\boldsymbol{x}} - \boldsymbol{x}\rangle \qquad (6.3.6)$$

和

$$\langle \boldsymbol{x}_\lambda^*, \bar{\boldsymbol{x}} - \boldsymbol{x}\rangle > \langle \boldsymbol{x}_2^*, \bar{\boldsymbol{x}} - \boldsymbol{x}\rangle. \qquad (6.3.7)$$

作 λ(6.3.6) + $(1-\lambda)$ × (6.3.7),则得到

$$\langle \boldsymbol{x}_\lambda^*, \bar{\boldsymbol{x}} - \boldsymbol{x}\rangle > \langle \boldsymbol{x}_\lambda^*, \bar{\boldsymbol{x}} - \boldsymbol{x}\rangle,$$

导致矛盾.▯

定理 6.3.2 设 $\boldsymbol{x} \in S$, $\boldsymbol{y} \in \psi(\boldsymbol{x})$, $\boldsymbol{p} \in \mathrm{int}K$, ψ 在 S 上是 K-凸的.

(1) 若 $\partial_w\psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}} \neq \varnothing$,则 $\partial_w\psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$ 是闭集.

(2) 若 $\psi(\boldsymbol{x})$ 是紧集,$\partial_w\psi(\boldsymbol{x})_{\boldsymbol{p}} \neq \varnothing$,则 $\partial_w\psi(\boldsymbol{x})_{\boldsymbol{p}}$ 是闭集.

证明 (1) 为证 $\partial_w\psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$ 是闭集,反之设点列 $\{\boldsymbol{x}_k^*\} \subset \partial_w\psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$, $\boldsymbol{x}_k^* \to \boldsymbol{x}^*$ $(k \to \infty)$,但 $\boldsymbol{x}^* \notin \partial_w\psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$,则从后者由定义 6.1.3 可知,存在 $\bar{\boldsymbol{x}} \in S$ 和 $\bar{\boldsymbol{y}} \in \psi(\bar{\boldsymbol{x}})$ 使得

$$\boldsymbol{y} - \bar{\boldsymbol{y}} + \langle \boldsymbol{x}^*, \bar{\boldsymbol{x}} - \boldsymbol{x}\rangle \boldsymbol{p} \in \mathrm{int}K.$$

因为 $\boldsymbol{y} - \bar{\boldsymbol{y}} + \langle \boldsymbol{x}_k^*, \bar{\boldsymbol{x}} - \boldsymbol{x}\rangle \boldsymbol{p} \to \boldsymbol{y} - \bar{\boldsymbol{y}} + \langle \boldsymbol{x}^*, \bar{\boldsymbol{x}} - \boldsymbol{x}\rangle \boldsymbol{p}$,所以存在充分大的 k' 使得

$$\boldsymbol{y} - \bar{\boldsymbol{y}} + \langle \boldsymbol{x}_{k'}^*, \bar{\boldsymbol{x}} - \boldsymbol{x}\rangle \boldsymbol{p} \in \mathrm{int}K.$$

由此,根据定义 6.1.3 知 $\boldsymbol{x}_{k'}^* \notin \partial_w\psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$,它与已设矛盾.

(2) 设点列 $\{\boldsymbol{x}_k^*\} \subset \partial_w\psi(\boldsymbol{x})_{\boldsymbol{p}}$, $\boldsymbol{x}_k^* \to \boldsymbol{x}^*$ $(k \to \infty)$,则存在 $\{\boldsymbol{y}^k\}$

$\subset \mathscr{Y}$, $y^k \in \psi(\bar{x})$，使 $x_k^* \in \partial_w \psi(x, y)_p$. 由于 $\psi(x)$ 是紧集，因而 $\{y^k\}$ 存在收敛的子列，不妨即设 $y^k \to y$ $(k \to \infty)$, $y \in \psi(x)$. 现假设 $x^* \notin \partial_w \psi(x, y)_p$，则由定义 6.1.3 的(2)可知，存在 $\bar{x} \in S$ 和 $\bar{y} \in \psi(\bar{x})$ 使得

$$y - \bar{y} + \langle x^*, \bar{x} - x \rangle p \in \text{int} K. \tag{6.3.8}$$

从 $x_k^* \to x^*$ 和 $y^k \to y$ 有

$$y^k - \bar{y} + \langle x_k^*, \bar{x} - x \rangle p \to y - \bar{y} + \langle x^*, \bar{x} - x \rangle p,$$

于是由(6.3.8)知，存在充分大的 k' 使

$$y^{k'} - \bar{y} + \langle x_{k'}^*, \bar{x} - x \rangle p \in \text{int} K.$$

据此，得到 $x_{k'}^* \notin \partial_w \psi(x, y_{k'})_p$，它与已设矛盾，从而有 $x^* \in \partial_w \psi(x, y)_p \subset \partial_w \psi(x)_p$. ▯

以下考虑锥弱次微分映射 $\partial_w \psi_p$: $S \to 2^{\mathscr{X}^*}$, $x \mapsto \partial_w \psi(x)_p$ 的上半连续性.

定理 6.3.3　设 $x \in S$, $y \in \psi(x)$, $p \in \text{int} K$，则集值映射 ψ: $S \to 2^{\mathscr{Y}}$ 在 (x, y) 处的 K-弱次微分映射 $\partial_w \psi_p$ 在点 (x, y) 处是上半连续的.

证明　设点列 $\{x^k\} \subset S$, $x^k \to x$ $(k \to \infty)$, $\{y^k\} \subset \mathscr{Y}$, $y^k \in \psi(x^k)$, $y^k \to y$ $(k \to \infty)$，有 $x_k^* \in \partial_w \psi(x^k, y^k)$ 和 $x_k^* \to x^*$ $(k \to \infty)$. 按定义 3.3.1 的(1)，反之假设 $x^* \notin \partial_w \psi(x, y)_p$，则由定义 6.1.3 的(2)可知存在 $\bar{x} \in S$, $\bar{y} \in \psi(\bar{x})$ 使得

$$y - \bar{y} + \langle x^*, \bar{x} - x \rangle p \in \text{int} K. \tag{6.3.9}$$

从 $x^k \to x$, $y^k \to y$ 和 $x_k^* \to x^*$ $(k \to \infty)$，有

$$y^k - \bar{y} + \langle x_k^*, \bar{x} - x^k \rangle p \to y - \bar{y} + \langle x^*, \bar{x} - x \rangle.$$

于是，由(6.3.9)得知，存在充分大的 k' 使

$$y^{k'} - \bar{y} + \langle x_{k'}^*, \bar{x} - x^{k'} \rangle p \in \text{int} K.$$

据此，得到 $x_{k'}^* \notin \partial_w \psi(x^{k'}, y^{k'})_p$，导致与已设矛盾. □

定义 6.3.1 设点 $x^0 \in S$，集合 $\psi(x) \subset \mathscr{Y}$ $(x \in S)$. 若存在点 x^0 的邻域 $U(x^0)$ 使 $\mathrm{cl} \bigcap_{x \in U(x^0)} \psi(x) \subset \mathscr{Y}$ 是紧集，则称 ψ 在点 x^0 附近是一致紧的.

定理 6.3.4 设 $x \in S$，$p \in \mathrm{int}K$. 若集值映射 $\psi: S \to 2^{\mathscr{Y}}$ 在 x 处是上半连续的，并且在 x 附近是一致紧的，则 K-弱次微分映射 $\partial_w \psi_p: S \to 2^{\mathscr{Y}^*}$，$x \mapsto \partial_w \psi(x)_p$ 在 x 处是上半连续的.

证明 设点列 $\{x^k\} \subset S$，$x^k \to x$ $(k \to \infty)$，$x_k^* \in \partial_w \psi(x^k)_p$ 和 $x_k^* \to x^*$ $(k \to \infty)$. 由定义 3.3.1 的(1)，只要证明 $x^* \in \partial_w \psi(x)_p$. 由于 $x_k^* \in \partial_w \psi(x^k)_p$，存在 $y^k \in \psi(x^k)$ 使得

$$x_k^* \in \partial_w \psi(x^k, y^k)_p.$$

因为 $x^k \to x$ $(k \to \infty)$ 和 ψ 在 x 附近是一致紧的，故知 $\{y^k\}$ 存在收敛子列. 不妨设 $y^k \to y$，由 ψ 在 x 处是上半连续的，按定义 3.3.1 的(1) 可知 $y \in \psi(x)$. 于是，根据定理 6.3.4 得到 $x^* \in \partial_w \psi(x, y)_p$，从而 $x^* \in \partial_w \psi(x)_p$. □

下面讨论锥弱次微分的连通性.

定义 6.3.2 设 S 是任意非空集合. 若不存在两个集合 $S_1 \neq \varnothing$ 和 $S_2 \neq \varnothing$，使得

$$S = S_1 \cup S_2,$$

其中 $\mathrm{cl}S_1 \cap S_2 = S_1 \cap \mathrm{cl}S_2 = \varnothing$，则称 S 是连通集，或 S 是连通的. 不是连通的集合称为是可分离的.

引理 6.3.5 设 $\mathscr{Y}$ 是线性空间. 若 $S \subset \mathscr{Y}$ 是非空凸集，则 S 是连通集.

证明 反之，假设 S 是可分离的，则由定义 6.3.2 可知，存在集合 $S_1 \neq \varnothing$ 和 $S_2 \neq \varnothing$，并且

$$\mathrm{cl}S_1 \cap S_2 = S_1 \cap \mathrm{cl}S_2 = \varnothing, \tag{6.3.10}$$

使得

$$S = S_1 \cup S_2. \tag{6.3.11}$$

设 $\boldsymbol{v}^1 \in S_1$, $\boldsymbol{v}^2 \in S_2$, 由(6.3.11)注意到 S 是凸集，故对任意的 $\lambda \in (0, 1)$ 有 $\lambda \boldsymbol{v}^1 + (1-\lambda)\boldsymbol{v}^2 \in S$，从而对于线段$[\boldsymbol{v}^1, \boldsymbol{v}^2]$有

$$[\boldsymbol{v}^1, \boldsymbol{v}^2] \subset S. \tag{6.3.12}$$

记 $S_3 = [\boldsymbol{v}^1, \boldsymbol{v}^2] \cap S_1$ 和 $S_4 = [\boldsymbol{v}^1, \boldsymbol{v}^2] \cap S_2$，因为 $\boldsymbol{v}^1 \in S_1$ 和 $\boldsymbol{v}^2 \in S_2$，故有 $S_3 \neq \varnothing$ 和 $S_4 \neq \varnothing$. 由(6.3.11)和(6.3.12)，我们有

$$\begin{aligned} S_3 \cup S_4 &= ([\boldsymbol{v}^1, \boldsymbol{v}^2] \cap S_1) \cup ([\boldsymbol{v}^1, \boldsymbol{v}^2] \cap S_2) \\ &= [\boldsymbol{v}^1, \boldsymbol{v}^2] \cap (S_1 \cup S_2) = [\boldsymbol{v}^1, \boldsymbol{v}^2] \cap S \\ &= [\boldsymbol{v}^1, \boldsymbol{v}^2]. \end{aligned}$$

又由(6.3.10)可知，有

$$\begin{aligned} \mathrm{cl}S_3 \cap S_4 &= \mathrm{cl}([\boldsymbol{v}^1, \boldsymbol{v}^2] \cap S_1) \cap ([\boldsymbol{v}^1, \boldsymbol{v}^2] \cap S_2) \\ &= [\boldsymbol{v}^1, \boldsymbol{v}^2] \cap \mathrm{cl}S_1 \cap S_2 = \varnothing. \end{aligned}$$

同理，有 $S_3 \cap \mathrm{cl}S_4 = \varnothing$. 由此，从以上所得根据定义 6.3.2 导致线段$[\boldsymbol{v}^1, \boldsymbol{v}^2]$是可分离的，但这是不可能的，从而定理得证. ▯

引理 6.3.6　设 $S \subset \mathscr{X}$ 是连通集，集值映射 $\psi: S \to 2^{\mathscr{Y}}$ 是上半连续的. 若对任意的 $\boldsymbol{x} \in S$, $\psi(\boldsymbol{x}) \neq \varnothing$ 是连通集，则 $\bigcup_{\boldsymbol{x} \in S} \psi(\boldsymbol{x})$ 是连通集.

证明　见文献[27]的命题 2.5.6. ▯

引理 6.3.7　设 $Y \subset \mathscr{Y}$ 是非空紧凸集，则 $\mathscr{E}_w(Y, K)$ 是连通集.

证明　见文献[27]的定理 3.5.12. ▯

定理 6.3.8　设 $\boldsymbol{x} \in S$, $\boldsymbol{y} \in \psi(\boldsymbol{x})$, $\psi: S \to 2^{\mathscr{Y}}$ 是集值映射，$\boldsymbol{p} \in \mathrm{int}K$.

(1) 若 $\partial_w \psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}} \neq \varnothing$，则 $\partial_w \psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$ 是连通集.

(2) 若 $\psi(\boldsymbol{x})$是紧凸集，ψ 在 S 上是上半连续的，则 $\partial_w \psi(\boldsymbol{x})_{\boldsymbol{p}}$ 是连通集.

证明　(1) 由定理 6.3.2 知 $\partial_w \psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$ 是凸集，再由引理

6.3.6 即知 $\partial_w\psi(\boldsymbol{x},\boldsymbol{y})_p$ 是连通集.

(2) 因 $\psi(\boldsymbol{x})$ 是紧集,故也是 K-半紧的. 由定理 6.2.10 可知,对任意的 $\boldsymbol{y}\in\mathscr{E}_w(\psi(\boldsymbol{x}),K)$ 有 $\partial_w\psi(\boldsymbol{x},\boldsymbol{y})_p\neq\varnothing$. 于是,据定理 6.1.1 的(3)和定义 6.1.3 的(2)得知

$$\begin{aligned}\partial_w\psi(\boldsymbol{x})_p &= \{\boldsymbol{x}^*\in\partial_w\psi(\boldsymbol{x},\boldsymbol{y})_p \mid \boldsymbol{y}\in\mathscr{E}_w(\psi(\boldsymbol{x}),K)\}\\ &= \bigcup_{\boldsymbol{y}\in\mathscr{E}_w(\psi(\boldsymbol{x}),K)}\partial_w\psi(\boldsymbol{x},\boldsymbol{y})_p.\end{aligned}\tag{6.3.13}$$

又因为 $\psi(\boldsymbol{x})$ 是紧凸集,故由定理 6.3.8 知 $\mathscr{E}_w(\psi(\boldsymbol{x}),K)$ 是连通集. 往下证明 K-弱次微分映射 $\partial_w\psi_p$ 关于 $\boldsymbol{y}$ 在 $\boldsymbol{y}$ 处是上半连续的. 事实上,设对任意的 $\boldsymbol{y}\in\mathscr{E}_w(\psi(\boldsymbol{x}),K)$,考虑点列

$$\{\boldsymbol{y}^k\}\subset\mathscr{Y},\ \boldsymbol{y}^k\to\boldsymbol{y}\ (k\to\infty),\ \boldsymbol{x}_k^*\in\partial_w\psi(\boldsymbol{x},\boldsymbol{y}^k)_p,\ \boldsymbol{x}_k^*\to\boldsymbol{x}^*.$$

根据定理 6.3.4 即知 $\partial_w\psi(\boldsymbol{x},\boldsymbol{z})_p$ 在 $\boldsymbol{y}$ 处是上半连续的. 最后,利用引理 6.3.7 和(6.3.13),便得 $\partial_w\psi(\boldsymbol{x})_p$ 是连通集. ▯

最后,讨论锥弱次微分的次单调性.

定理 6.3.9 设 $\boldsymbol{x}^1,\boldsymbol{x}^2\in S$, $\boldsymbol{y}^1\in\psi(\boldsymbol{x}^1)$, $\boldsymbol{y}^2\in\psi(\boldsymbol{x}^2)$. 若 $\boldsymbol{x}_1^*\in\partial_w\psi(\boldsymbol{x}^1,\boldsymbol{y}^1)_p$ 和 $\boldsymbol{x}_2^*\in\partial_w\psi(\boldsymbol{x}^2,\boldsymbol{y}^2)_p$,则 $\langle\boldsymbol{x}_1^*-\boldsymbol{x}_2^*,\boldsymbol{x}^1-\boldsymbol{x}^2\rangle\geqslant 0$.

证明 若 $\boldsymbol{x}_1^*=\boldsymbol{x}_2^*$ 或者 $\boldsymbol{x}^1=\boldsymbol{x}^2$,则定理结论显然成立. 以下设 $\boldsymbol{x}_1^*\neq\boldsymbol{x}_2^*$, $\boldsymbol{x}^1\neq\boldsymbol{x}^2$. 用反证法,假设 $\delta=\langle\boldsymbol{x}_1^*-\boldsymbol{x}_2^*,\boldsymbol{x}^1-\boldsymbol{x}^2\rangle<0$. 记

$$\boldsymbol{a}^1=\boldsymbol{y}^1-\boldsymbol{y}^2-\langle\boldsymbol{x}_1^*,\boldsymbol{x}^1-\boldsymbol{x}^2\rangle\boldsymbol{p},$$

$$\boldsymbol{a}^2=\boldsymbol{y}^2-\boldsymbol{y}^1-\langle\boldsymbol{x}_2^*,\boldsymbol{x}^2-\boldsymbol{x}^1\rangle\boldsymbol{p}.$$

对任意的 $\lambda\in[0,1]$,令 $F(\lambda)=\lambda\boldsymbol{a}^1+(1-\lambda)\boldsymbol{a}^2$,则有

$$F(\lambda)=(2\lambda-1)\boldsymbol{a}^1-(1-\lambda)\delta\boldsymbol{p},\tag{6.3.14}$$

或即

$$F(\lambda)=(1-2\lambda)\boldsymbol{a}^2-\lambda\delta\boldsymbol{p}.\tag{6.3.15}$$

作集合

$$E=\{\boldsymbol{y}\in\mathscr{Y}\mid\boldsymbol{y}=F(\lambda)+(1/2)\delta\boldsymbol{p},\ \lambda\in[0,1]\}.$$

显然 E 是凸集，往下证明

$$E \cap \mathrm{int}K = \varnothing. \tag{6.3.16}$$

事实上，反之，若存在 $\boldsymbol{q} \in E \cap \mathrm{int}K$，由 E 的定义有

$$F(\lambda) + (1/2)\delta\boldsymbol{p} = \boldsymbol{q}, \text{对某个 } \lambda \in [0, 1]. \tag{6.3.17}$$

当 $\lambda = 1/2$ 时，得 $\boldsymbol{q} = \boldsymbol{0} \in \mathrm{int}K$，导致矛盾. 当 $0 \leqslant \lambda < 1/2$ 时，由(6.3.15)和(6.3.17)得 $\boldsymbol{a}^2 = \dfrac{2\boldsymbol{q} - (2\lambda - 1)\delta\boldsymbol{p}}{2(1 - 2\lambda)} \in \mathrm{int}K$，即

$$\boldsymbol{y}^2 - \boldsymbol{y}^1 - \langle \boldsymbol{x}_2^*, \boldsymbol{x}^2 - \boldsymbol{x}^1 \rangle \boldsymbol{p} \in \mathrm{int}K.$$

由此，根据定义 6.1.3 的(2)知 $\boldsymbol{x}_2^* \notin \partial_w \psi(\boldsymbol{x}^2, \boldsymbol{y}^2)_p$，这与已设矛盾. 当 $1/2 < \lambda \leqslant 1$ 时，由(6.3.14)和(6.3.17)得 $\boldsymbol{a}^1 = \dfrac{2\boldsymbol{q} - (1 - 2\lambda)\delta\boldsymbol{p}}{2(2\lambda - 1)} \in \mathrm{int}K$，即

$$\boldsymbol{y}^1 - \boldsymbol{y}^2 - \langle \boldsymbol{x}_1^*, \boldsymbol{x}^1 - \boldsymbol{x}^2 \rangle \boldsymbol{p} \in \mathrm{int}K.$$

于是，由定义 6.1.3 的(2)知 $\boldsymbol{x}_1^* \notin \partial_w \psi(\boldsymbol{x}^1, \boldsymbol{y}^1)_p$，也与已设矛盾.

现在，从(6.3.16)由定理 1.3.7 的(1)可知，存在 $\boldsymbol{y}^* \in \mathscr{Y}^* \setminus \{\boldsymbol{0}\}$ 使得

$$\langle \boldsymbol{y}^*, \boldsymbol{q} \rangle \geqslant \langle \boldsymbol{y}^*, \boldsymbol{y} \rangle \quad \forall \boldsymbol{q} \in K, \forall \boldsymbol{y} \in E. \tag{6.3.18}$$

因为 $\boldsymbol{0} \in E$，所以从(6.3.18)得 $\langle \boldsymbol{y}^*, \boldsymbol{q} \rangle \geqslant 0\ (\forall \boldsymbol{q} \in K)$，从而 $\boldsymbol{y}^* \in K^* \setminus \{\boldsymbol{0}\}$. 由 $\boldsymbol{p} \in \mathrm{int}K$ 和定理 1.4.11 的(1)得到 $\langle \boldsymbol{y}^*, \boldsymbol{p} \rangle > 0$. 将 $\boldsymbol{a}^1 + (1/2)\delta\boldsymbol{p}$，$\boldsymbol{a}^2 + (1/2)\delta\boldsymbol{p} \in E$ 代入(6.3.18)，我们有

$$\langle \boldsymbol{y}^*, \boldsymbol{a}^1 \rangle + (1/2)\delta\boldsymbol{p} \leqslant 0, \tag{6.3.19}$$

$$\langle \boldsymbol{y}^*, \boldsymbol{a}^2 \rangle + (1/2)\delta\boldsymbol{p} \leqslant 0. \tag{6.3.20}$$

由(6.3.18)+(6.3.19)得 $\langle \boldsymbol{y}^*, -\delta\boldsymbol{p} \rangle + (1/2)\delta\boldsymbol{p} \leqslant 0$，即 $\delta\langle \boldsymbol{y}^*, \boldsymbol{p} \rangle \geqslant 0$，因此推知 $\delta \geqslant 0$，导致与假设矛盾. ▯

定理 6.3.10 设 $\boldsymbol{x} \in S$，$\boldsymbol{y} \in \psi(\boldsymbol{x})$，$\boldsymbol{p}^1$，$\boldsymbol{p}^2 \in \mathrm{int}K$. 若 $\boldsymbol{x}^* \in \partial_w \psi(\boldsymbol{x}, \boldsymbol{y})_{p^1}$，则存在 $\boldsymbol{q}^* \in K^* \setminus \{\boldsymbol{0}\}$ 使得 $\alpha \boldsymbol{x}^* \in \partial_w \psi(\boldsymbol{x}, \boldsymbol{y})_{p^2}$，其中

$\alpha = \dfrac{\langle q^*, p^1\rangle}{\langle q^*, p^2\rangle}$.

证明 由定理 6.2.12 知，存在 $q^* \in K^*\setminus\{\boldsymbol{0}\}$ 使得

$$\langle q^*, y' - y\rangle \geqslant \langle x^*, x' - x\rangle\langle q^*, p^1\rangle$$

$$\forall x' \in S, \forall y' \in \psi(x').$$

令 $\alpha = \dfrac{\langle q^*, p^1\rangle}{\langle q^*, p^2\rangle}$，注意 $\langle q^*, p^1\rangle > 0$ 和 $\langle q^*, p^2\rangle > 0$ 有 $\alpha > 0$，故得

$$\langle q^*, y' - y\rangle \geqslant \langle \alpha x^*, x' - x\rangle\langle q^*, p^2\rangle$$

$$\forall x' \in S, \forall y' \in \psi(x').$$

由定理 6.2.11 的(2)，即得 $\alpha x^* \in \partial_w\psi(x, y)_{p^2}$. ▯

注 6.3.1 定理 6.3.11 表明了 $\partial_w\psi(x, y)_{p^1}$ 与 $\partial_w\psi(x, y)_{p^2}$ 中的元素仅相差一个倍数. 因此，只要求得关于某个 p 的锥弱次微分就可得到所有的结果. 从这一意义上，我们说 $\partial_w\psi(x, y)_p$ 与 p 是无关的.

定理 6.3.11 设实值向量函数 $f: S \to R^m$ 是 R^m_+-凸的，$f(x) = (f_1(x), \cdots, f_m(x))^T$ 在 S 上每一点处有定义，令 $p = (1, \cdots, 1)^T \in R^m$. 若 f 在 S 上是连续的，$x \in S$，则

$$\partial_w f(x)_p = \left\{ \sum_{i=1}^m \mu_i \partial f_i(x) \,\middle|\, \sum_{i=1}^m \mu_i = 1, \mu_i \geqslant 0, 1, \cdots, m \right\},$$

其中 $\partial f_i(x)$ 是凸函数的次微分.

证明 由定理 6.2.13 的(2)和凸函数次微分的可加性，可得

$$\partial_w f(x)_p = \bigcup_{q^* \in K^*\setminus\{\boldsymbol{0}\}} \frac{1}{\langle q^*, p\rangle} \partial\theta(x, \langle q^*, y\rangle)_1$$

$$= \bigcup_{q^* \in K^*\setminus\{\boldsymbol{0}\}} \left(\sum_{i=1}^m q_i^* \right)^{-1} \sum_{i=1}^m q_i^* \partial f_i(x)$$

$$= \left\{ \sum_{i=1}^{m} \mu_i \partial f_i(x) \,\middle|\, \sum_{i=1}^{m} \mu_i = 1,\ \mu_i \geqslant 0,\ i = 1, \cdots, m \right\}.$$

于是,定理得证. ▯

显然,例 6.2.2 可看成是定理 6.3.12 的推论.

例 6.3.1　考虑实值向量函数 $\boldsymbol{f}: R^2 \to R^2$, $\boldsymbol{f}(x, y) = (|x|, |y|)^T$, $(x, y)^T \in R^2$. 设 $f_i(x) = |x|$, $x \in R$ $(i = 1, 2)$, 则有 $\partial f_i(0) = [-1, 1]$,根据定理 6.3.12 得

$$\partial_w \boldsymbol{f}(0, 0)_p = \{(\mu_1 a, \mu_2 b)^T \in R^2 \mid \mu_1 + \mu_2 = 1,$$
$$\mu_1, \mu_2 \geqslant 0, a, b \in [-1, 1]\}.$$

此外,当 $x > 0$, $y = 0$ 时,有

$$\partial_w \boldsymbol{f}(x, 0)_p = \{(\mu_1, \mu_2 b)^T \in R^2 \mid \mu_1 + \mu_2 = 1,$$
$$\mu_1, \mu_2 \geqslant 0, b \in [-1, 1]\};$$

当 $x = 0$, $y > 0$ 时,有

$$\partial_w \boldsymbol{f}(0, y)_p = \{(\mu_1 a, \mu_2)^T \in R^2 \mid \mu_1 + \mu_2 = 1,$$
$$\mu_1, \mu_2 \geqslant 0, a \in [-1, 1]\};$$

当 $x > 0$, $y > 0$ 时,有

$$\partial_w \boldsymbol{f}(x, y)_p = \{(\mu_1, \mu_2)^T \in R^2 \mid \mu_1 + \mu_2 = 1, \mu_1, \mu_2 \geqslant 0\};$$

当 $x > 0$, $y < 0$ 时,有

$$\partial_w \boldsymbol{f}(x, y)_p = \{(\mu_1, -\mu_2)^T \in R^2 \mid \mu_1 + \mu_2 = 1, \mu_1, \mu_2 \geqslant 0\};$$

当 $x < 0$, $y > 0$ 时,有

$$\partial_w \boldsymbol{f}(x, y)_p = \{(-\mu_1, \mu_2)^T \in R^2 \mid \mu_1 + \mu_2 = 1, \mu_1, \mu_2 \geqslant 0\};$$

当 $x < 0$, $y < 0$ 时,有

$$\partial_w f(x, y)_p = \{(-\mu_1, -\mu_2)^T \in R^2 \mid \mu_1 + \mu_2 = 1, \mu_1, \mu_2 \geq 0\}.$$

$\partial_w f(x, y)_p$ 的 9 种情况如表 6.3.1 所示.

表 6.3.1

x	$x=0$			$x>0$			$x<0$		
y	$y=0$	$y>0$	$y<0$	$y=0$	$y>0$	$y<0$	$y=0$	$y>0$	$y<0$
$\partial_w f(x,y)_p$	$ABCD$	ABC	ACD	BCD	BC	CD	ABD	AB	AD

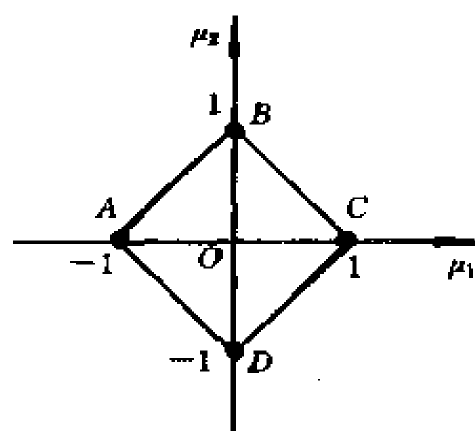


图 6.3.1

图 6.3.1 中的 $ABCD$ 表示由直线 AB、BC、CD 和 AD 围成的方形区域；ABC 表示由直线 AB、BC 和 CA 围成的三角形区域；AB 表示直线. $\partial_w f(x, y)_p$ 与相应的区域对应.

§6.4 运算性质

本节给出锥弱次微分的基本运算性质，如次可加性和复合映射的链式法则等.

设集合 $S \subset \mathscr{X}$, $\psi: \mathscr{X} \to 2^{\mathscr{Y}}$, $\boldsymbol{x} \mapsto \psi(\boldsymbol{x})$ 是集值映射, $K \subset \mathscr{Y}$ 是内部非空的尖闭凸锥.

定理 6.4.1 设 $S \subset \mathscr{X}$ 是非空凸集，集值映射 $\psi: S \to 2^{\mathscr{Y}}$ 是 K-凸的，$\boldsymbol{x} \in S$, $\boldsymbol{y} \in \psi(\boldsymbol{x})$, $\boldsymbol{p} \in \mathrm{int}K$. 若 $\alpha > 0$，则

$$\partial_w(\alpha\psi)(\boldsymbol{x}, \boldsymbol{y})_p = \alpha\partial_w\psi(\boldsymbol{x}, \boldsymbol{y})_p. \tag{6.4.1}$$

证明　设 $\alpha = 0$，则(6.4.1)显然成立. 设 $\alpha > 0$，任取 $x^* \in \partial_w(\alpha\psi)(x, y)_p$，由定理 6.2.12 知存在 $q^* \in K^*\setminus\{0\}$ 使得

$$\langle q^*, \alpha y' - \alpha y\rangle \geqslant \langle x^*, x' - x\rangle\langle q^*, p\rangle$$

$$\forall x' \in S, \forall \alpha y' \in \alpha\psi(x'),$$

即

$$\langle q^*, y' - y\rangle \geqslant \langle x^*/\alpha, x' - x\rangle\langle q^*, p\rangle$$

$$\forall x' \in S, \forall y' \in \psi(x').$$

由此，根据定理 6.2.11 的(2)得知 $x^* \in \alpha\partial_w\psi(x, y)_p$. 推导的逆过程也是成立的. 定理证毕. □

定理 6.4.2　设 $S \subset \mathscr{X}$ 是内部非空的凸集，集合 $\psi_1(x)$，$\psi_2(x) \subset \mathscr{Y}$ $(x \in S)$ 非空，

$$(\psi_1 + \psi_2)(x) = \{y^1 + y^2 \mid y^1 \in \psi_1(x), y^2 \in \psi_2(x), x \in S\},$$

$$x \in \operatorname{int}S, y^1 \in \psi_1(x), y^2 \in \psi_2(x).$$

对于集值映射 $\psi_1: S \to 2^{\mathscr{Y}}, x \mapsto \psi_1(x)$；$\psi_2: S \to 2^{\mathscr{Y}}, x \mapsto \psi_2(x)$；以及 $\psi_1 + \psi_2: S \to 2^{\mathscr{Y}}, x \mapsto (\psi_1 + \psi_2)(x)$，若 ψ_1 和 ψ_2 在 S 上是 K-凸的，并且 ψ_1 或 ψ_2 在点 x 处是 K-有界逆开的(或连通的)，则

$$\partial_w(\psi_1 + \psi_2)(x, y^1 + y^2)_p \subset \partial_w\psi_1(x, y^1)_p + \partial_w\psi_2(x, y^2)_p, \tag{6.4.2}$$

$$\partial_w(\psi_1 + \psi_2)(x)_p \subset \partial_w\psi_1(x)_p + \partial_w\psi_2(x)_p. \tag{6.4.3}$$

证明　设 $x^* \in \partial_w(\psi_1 + \psi_2)(x, y)_p$，不妨设 ψ_2 在 x 处满足 K-有界逆开的. 作集合

$$E = \{(x', y') \in \mathscr{X} \times \mathscr{Y} \mid x' \in S, y' = y^1 - y_1'$$

$$+ \langle x^*, x' - x\rangle p - q^1, y_1' \in \psi_1(x'), q^1 \in K\},$$

$$F = \{(x', y') \in \mathscr{X} \times \mathscr{Y} \mid x' \in S, y' = y_2' - y^2 + q^2,$$

$$y_2' \in \psi_2(x');\ q^2 \in K\}.$$

显然 $E \neq \varnothing$；根据引理 6.2.1 知 $\mathrm{int}(K\text{-epi}\psi_2) \neq \varnothing$，从而 $\mathrm{int}F \neq \varnothing$. 由已知条件不难推得 E 和 F 是凸集，下面证明

$$E \cap \mathrm{int}F = \varnothing. \tag{6.4.4}$$

事实上，否则，假若 $E \cap \mathrm{int}F \neq \varnothing$，则存在 $\bar{x} \in S$，$\bar{y}_1 \in \psi_1(\bar{x})$，$\bar{y}_2 \in \psi_2(\bar{x})$，以及 $\bar{q}_1$，$\bar{q}_2 \in K$，使得

$$\begin{aligned}\bar{y} &= \bar{y}_2 - y^2 + \bar{q}_2 \\ &= y^1 - \bar{y}_1 + \langle x^*, \bar{x} - x\rangle p - \bar{q}_1.\end{aligned} \tag{6.4.5}$$

由于 $(\bar{x}, \bar{y}) \in \mathrm{int}F$，因而存在充分小的 $\varepsilon > 0$ 有 $(\bar{x}, \bar{y} - \varepsilon p) \in F$. 据此，有 $y_2' \in \psi_2(\bar{x})$，$q_2' \in K$，使得

$$y_2' - y^2 + q_2' = \bar{y}_2 - y^2 + \bar{q}_2 - \varepsilon p.$$

将上式代入(6.4.5)，有

$$y_2' - y^2 + q_2' + \varepsilon p = y^1 - \bar{y}_1 + \langle x^*, \bar{x} - x\rangle p - \bar{q}_1,$$

因为 K 是凸锥，故

$$\begin{aligned}y^1 + y^2 - \langle x^*, x\rangle p - (\bar{y}_1 + y_2' - \langle x^*, \bar{x}\rangle p) \\ = \bar{q}_1' + \varepsilon p \in \mathrm{int}K.\end{aligned}$$

于是，由定义 6.1.3 的(2)知 $x^* \notin \partial_w(\psi_1 + \psi_2)(x, y)_p$，这与已设矛盾.

由(6.4.4)根据定理 1.3.7 的(1)，得知存在 $(x^*, y^*) \in \mathscr{X}^* \times \mathscr{Y}^*$ 和 $(x^*, y^*) \neq \mathbf{0}$，使得

$$\langle x^*, x_e\rangle + \langle y^*, y_e\rangle \geqslant \beta \geqslant \langle x^*, x_f\rangle + \langle y^*, y_f\rangle$$

$$\forall (x_e, y_e) \in E \quad \forall (x_f, y_f) \in F. \tag{6.4.6}$$

在上式中 $y^* \neq \mathbf{0}$，因为否则，令 $x_e = x'$，$x_f = x$ 代入(6.4.6)得 $\langle x^*, x' - x\rangle \geqslant 0$ $(\forall x' \in S)$. 类似于定理 6.2.4 的证明，导致与

$x \in \operatorname{int}S$ 矛盾. 由于对任意的 $q \in K$ 有 $(x, -q) \in E$ 和 $(x, q) \in F$,代入(6.4.6)得

$$\langle -y^*, q\rangle > 0 \quad \forall q \in K,$$

因此 $-y^* \in K^*\setminus\{0\}$. 令 $q^* = -y^*$, 由 $p \in \operatorname{int}K$, 根据定理 1.4.11 的(1)可知 $\langle q^*, p\rangle > 0$. 又由 $(x, 0) \in E \cap F$, 代入(6.4.6)得 $\beta = \langle x^*, x\rangle$. 设

$$x_e = x_f = x', \ y_e = \bar{y}_1 - y_1' + \langle x^*, x' - x\rangle p,$$

$$y_f = y_2' - y^2,$$

将以上式子一同代入(6.4.6),得到

$$\langle x^*, x'\rangle + \langle -q^*, y^1 - y_1'\rangle + \langle x^*, x' - x\rangle p$$

$$\geqslant \langle x^*, x\rangle \geqslant \langle x^*, x'\rangle + \langle -q^*, y_2' - y^2\rangle.$$

令 $x_2^* = x^*/\langle q^*, p\rangle$, $x_1^* = x^* - x_2^*$, 则有

$$\langle q^*, y_1' - y^1\rangle \geqslant \langle x_1^*, x' - x\rangle\langle q^*, p\rangle$$

$$\forall x' \in S, \forall y_1' \in \psi_1(x'),$$

$$\langle q^*, y_2' - y^2\rangle \geqslant \langle x_2^*, x' - x\rangle\langle q^*, p\rangle$$

$$\forall x' \in S, \forall y_2' \in \psi_2(x').$$

据此,由定理 6.2.11 的(2)知 $x_1^* \in \partial\psi_1(x, y^1)_p$, $x_2^* \in \partial\psi_2(x, y^2)_p$, 因而(6.4.2)成立.

由(6.4.2)和定义 6.1.3 的(2),容易推出(6.4.3)成立. ▯

注 6.4.1　仅当 ψ_1 和 ψ_2 是单值凸函数时,(6.4.2)才成为等式.

注 6.4.2　(6.4.2)的反包含关系一般不成立,即

$$\partial_w\psi_1(x, y^1)_p + \partial_w\psi_2(x, y^2)_p \subset \partial_w(\psi_1 + \psi_2)(x, y^1 + y^2)_p$$

不一定成立. 例如,设 $\mathscr{X} = R$, $\mathscr{Y} = R^2$, $K = R_+^2$, $p = (1, 1)^T$,

$\psi_1(x) = (-x, x)^T$, $\psi_2(x) = (x, -x)^T$,则 $\psi_1, \psi_2: R \to R^2$ 是连续的和 R^2_+ - 凸的,并且 $\psi_1(0) = \boldsymbol{0}$, $\psi_2(0) = \boldsymbol{0}$. 由定理 6.3.12 得

$$\partial_w \psi_1(0)_p = \{-\mu_1 + \mu_2 \mid \mu_1 + \mu_2 = 1, \mu_1, \mu_2 \geqslant 0\},$$

$$\partial_w \psi_2(0)_p = \{\mu_1 - \mu_2 \mid \mu_1 + \mu_2 = 1, \mu_1, \mu_2 \geqslant 0\},$$

故有

$$\partial_w \psi_1(0)_p + \partial_w \psi_2(0)_p = \{2a - 2b \mid 0 \leqslant a \leqslant 1, 0 \leqslant b \leqslant 1\}.$$

另外,由于 $(\psi_1 + \psi_2)(x) = (0, 0)^T \ (\forall x \in R)$,则有

$$\partial_w (\psi_1 + \psi_2)(0)_p = \{\boldsymbol{0}\}.$$

推论 6.4.3 设 $S \subset \mathscr{X}$ 是内部非空的开凸集,集合 $\Psi(\boldsymbol{x}) \subset R^m (\boldsymbol{x} \in S, m \geqslant 1)$ 是非空的,$\boldsymbol{p} = (1, \cdots, 1)^T \in R^m$. 记

$$\Psi(\boldsymbol{x}) = (\psi_1(\boldsymbol{x}), \cdots, \psi_m(\boldsymbol{x}))^T \subset R^m,$$

若集值映射 $\Psi: S \to 2^{R^m}$ 是 R^m_+ - 凸的,$\boldsymbol{x} \in \mathrm{int}S$, $\Psi(\boldsymbol{x})$ 在 $\boldsymbol{x}$ 处是连通的,则

$$\partial_w \Psi(\boldsymbol{x})_p \subset \left\{ \sum_{i=1}^m \mu_i \partial_w \psi_i(\boldsymbol{x}) \,\middle|\, \sum_{i=1}^m \mu_i = 1, \right.$$

$$\left. \mu_i \in [0, 1, i = 1, \cdots, m \right\}. \qquad (6.4.7)$$

当 Ψ 是单值函数时,上式成为等式.

证明 若 $\partial_w \Psi(\boldsymbol{x})_p = \varnothing$,则(6.4.7)显然成立. 设 $\partial_w \Psi(\boldsymbol{x})_p \neq \varnothing$,则存在 $\boldsymbol{y} \in \Psi(\boldsymbol{x})$,使得 $\boldsymbol{x}^* \in \partial_w \Psi(\boldsymbol{x}, \boldsymbol{y})_p$. 由此,根据定理 6.2.12,可知存在 $\boldsymbol{q}^* \in R^m_+ \setminus \{\boldsymbol{0}\}$ 使得

$$\langle \boldsymbol{q}^*, \boldsymbol{y}' - \boldsymbol{y} \rangle \geqslant \langle \boldsymbol{x}^*, \boldsymbol{x}' - \boldsymbol{x} \rangle \langle \boldsymbol{q}^*, \boldsymbol{p} \rangle$$

$$\forall \boldsymbol{x}' \in S, \forall \boldsymbol{y}' \in \Psi(\boldsymbol{x}'). \qquad (6.4.8)$$

设 $\boldsymbol{q}^* = (q_1^*, \cdots, q_m^*)^T$,记 $\alpha = \langle \boldsymbol{q}^*, \boldsymbol{p} \rangle = \sum_{i=1}^m q_i^*$,作集值函数

$$F: S \to 2^R,\ F(x) = \{f \mid f = \langle q^*,\ y\rangle,\ y \in \psi(x)\},$$

由(6.4.8)得

$$f - f' \geqslant \langle \alpha x^*,\ x' - x\rangle,\quad x' \in S,\ f \in F(x'),$$

其中 $f' = \langle q^*,\ y\rangle$. 因为 ψ 是 R_+^m - 凸的，则 F 是 R_+^m - 凸的，由定理 6.2.11 得 $\alpha x^* \in \partial_w F(x,\ f')_1$. 再由 ψ 的分量 $\psi_i(i = 1,\ \cdots,\ m)$ 在 S 上是 R_+ - 凸的，并且在 x 处是连通的，根据定理 6.4.1 和定理 6.4.2，有

$$\partial_w F(x)_1 \subset \sum_{i=1}^m q_i^* \partial_w \psi_i(x)_1,$$

因此

$$x^* \in \sum_{i=1}^m \frac{q_i^*}{\alpha} \partial_w \psi_i(x)_1.$$

令 $\mu_i = q_i^* / \alpha(i = 1,\ \cdots,\ m)$，即得(6.4.7).

当 ψ 是单值函数时，由定理 6.3.12 即知结论成立. □

最后，讨论复合映射的求锥弱次微分的问题.

定义 6.4.1　设 $S \subset \mathscr{X}$ 是非空凸集，$\psi: S \to 2^{\mathscr{Y}}$ 是集值映射. 若对任意的 $\lambda \in (0,\ 1)$ 有

$$\lambda\psi(x^1) + (1 - \lambda)\psi(x^2)$$

$$\subset \psi(\lambda x^1 + (1 - \lambda)x^2)\quad \forall x^1,\ x^2 \in S,$$

则称 ψ 是 S 上的凸集值映射，或 ψ 在 S 上是凸的.

显然，集值映射 ψ 在 S 上是凸的，则对任何尖闭凸锥 $K \subset \mathscr{Y}$，它在 S 上也是 K- 凸的.

设 $\mathscr{X}$，$\mathscr{Y}$ 和 $\mathscr{Z}$ 是局部凸线性拓扑空间，$K_1 \subset \mathscr{Z}$ 是内部非空的尖闭凸锥.

定理 6.4.4　设 $S \subset \mathscr{X}$ 是非空凸集，集合 $F(x) \subset \mathscr{Y}$ $(\forall x \in S)$ 是非空的，$Y = F(S) = \bigcup_{x \in S} F(x)$，集合 $G(y) \subset \mathscr{Z}$ $(y \in Y)$ 是非空的. 又设集值映射 $F: S \to 2^{\mathscr{Y}}$ 在 S 上是凸的，集值映射 $G: Y \to$

$2^{\mathscr{Z}}$ 在 Y 上是 K_1- 凸的，$p_1 \in \text{int}K_1$. 考虑集值映射

$$G \circ F: S \to 2^{\mathscr{Z}},\ x \mapsto G \circ F(x) = \bigcup_{y \in F(x)} G(y),$$

设 $x \in S$, $y \in F(x)$, $z \in G(y)$, $y \in \text{int}Y$. 若 G 在 y 处是 K_1- 有界逆开的或连通的，则

$$\partial_w G \circ F(x, z)_{p_1} = \left\{ \partial\theta(x, \langle y^*, y \rangle)_1 \,\middle|\, y^* \in \partial_w G(y, z)_{p_1} \right\}, \tag{6.4.9}$$

其中集值函数

$$\theta: S \to 2^R,\ x \mapsto \theta(x) = \left\{ \langle y^*, y \rangle \,|\, y \in F(x) \right\},$$

$\partial\theta(x, \cdot)_1$ 是 θ 在 x 处关于 $p = 1$ 的 R_+ 一次微分.

证明 设 $x^* \in \left\{ \partial\theta(x, \langle y^*, y \rangle)_1 \,|\, y^* \in \partial_w G(y, z)_{p_1} \right\}$，先证明 θ 在 S 上是 R_+- 凸的. 为此对任意的 x^1, $x^2 \in S$ 和任意的 $\lambda \in (0, 1)$，任取 $y^1 \in F(x^1)$, $y^2 \in F(x^2)$，由 F 在 S 上是凸的，则

$$\lambda y^1 + (1 - \lambda) y^2 \in F\left(\lambda x^1 + (1 - \lambda) x^2 \right).$$

因为 $\lambda \langle y^*, y^1 \rangle + (1 - \lambda) \langle y^*, y^2 \rangle = \langle y^*, \lambda y^1 + (1 - \lambda) y^2 \rangle$，所以有

$$\langle y^*, \lambda y^1 + (1 - \lambda) y^2 \rangle \in \langle y^*, F(\lambda x^1 + (1 - \lambda) x^2) \rangle.$$

由此有 $\lambda\theta(x^1) + (1 - \lambda)\theta(x^2) \subset \theta(\lambda x^1 + (1 - \lambda) x^2)$. 再由 F 在 S 上是凸的，得知 Y 是凸集. 下面证明 $G \circ F$ 在 S 上是 K_1- 凸的. 设对任意的 x^1, $x^2 \in S$ 和任意的 $\lambda \in (0, 1)$，任取 $z^1 \in G \circ F(x^1)$, $z^2 \in G \circ F(x^2)$，则存在 $y^1 \in F(x^1)$ 和 $y^2 \in F(x^2)$，使得 $z^1 \in G(y^1)$ 和 $z^2 \in G(y^2)$. 由于 G 在 Y 上是 K_1- 凸的，我们有

$$\lambda z^1 + (1 - \lambda) z^2 \in G\left(\lambda y^1 + (1 - \lambda) y^2 \right) + K_1. \tag{6.4.10}$$

另外，由 F 在 S 上是凸的，则

$$\lambda y^1 + (1-\lambda)y^2 \in F\left(\lambda x^1 + (1-\lambda)x^2\right).$$

根据(6.4.10)和上式，有

$$\lambda z^1 + (1-\lambda)z^2 \in G\left(F(\lambda x^1 + (1-\lambda)x^2\right) + K_1.$$

由此得知 θ 在 S 上是 R_+-凸的，$G \circ F$ 在 S 上是 K_1-凸的. 因为

$$x^* \in \partial\theta\left(x, \langle y^*, y\rangle\right)_1,$$

由定理 6.2.12 可得

$$\langle y^*, y'\rangle - \langle y^*, y\rangle \geqslant \langle x^*, x' - x\rangle$$

$$\forall x' \in S, \forall y' \in F(x'). \tag{6.4.11}$$

再根据 $y^* \in \partial_w G(y, z)_{p_1}$，由定理 6.2.12 可知，存在 $q_1^* \in K_1^* \setminus \{\theta\}$ 使得

$$\langle q_1^*, z' - z\rangle \geqslant \langle y^*, y' - y\rangle\langle q_1^*, p_1\rangle$$

$$\forall y' \in Y, \forall z' \in G(y').$$

从上式和(6.4.11)，我们有

$$\langle q_1^*, z' - z\rangle \geqslant \langle x^*, x' - x\rangle\langle q_1^*, p_1\rangle$$

$$\forall x' \in S, \forall z' \in G \circ F(x').$$

根据定理 6.2.11 的(2)，便得 $x^* \in \partial_w G \circ F(x, z)_{p_1}$.

现设 $x^* \in \partial_w G \circ F(x, z)_{p_1}$，作集合

$$S_1 = \{(y', z') \in \mathscr{Y} \times \mathscr{Z} \mid y' \in Y, z' = \bar{z} + q_1' - z,$$

$$\bar{z} \in G(y'), q_1' \in K_1\},$$

$$S_2 = \{(y', z') \in \mathscr{Y} \times \mathscr{Z} \mid x' \in S, y' \in F(x'),$$

$$z' = \langle x^*, x' - x\rangle p_1\}.$$

由于 S 和 Y 是凸的，F 在 S 上是凸的，G 在 Y 上是 K_1-凸的，容易

推知 S_1 和 S_2 是 $\mathscr{Y}\times\mathscr{Z}$ 中的凸集. 由已知 G 在 y 处是 K_1-有界逆开的，根据引理 6.2.1 得 $\mathrm{int}S_1\neq\varnothing$. 下面证明

$$S_2\cap \mathrm{int}S_1=\varnothing. \qquad (6.4.12)$$

事实上，否则，假若存在 $\bar{x}\in S$，$\bar{y}\in F(\bar{x})$，$\bar{z}\in G(\bar{y})$ 和 $q_1'\in K_1$，使得

$$y'=\bar{y},\ z'=\langle x^*,\bar{x}-x\rangle p_1=\bar{z}+q_1'-z, \quad (6.4.13)$$

并且 $(y',z')\in\mathrm{int}S_1$，则存在正数 $\lambda>0$ 使得 $(y',z'-\lambda p_1)\in S_1$，故又存在 $\bar{z}'\in G(y')$ 和 $\bar{q}'_1\in K_1$，使得

$$z'-\lambda p_1=\bar{z}'+\bar{q}_1'-z.$$

由(6.4.13)和上式可知 $\bar{z}'\in G\circ F(\bar{x})$ 和

$$\bar{z}'+\bar{q}_1'-z+\lambda p_1=\langle x^*,\bar{x}-x\rangle p_1.$$

整理上式，得

$$z-\langle x^*,x\rangle p_1-\big(\bar{z}'-\langle x^*,\bar{x}\rangle p_1\big)=\lambda p_1+\bar{q}_1'\in\mathrm{int}K_1.$$

由此，按定义 6.1.3 的(2)得到 $x^*\notin\partial_w G\circ F(x,z)_{p_1}$，这与已设矛盾. 由(6.4.12)，根据定理 1.3.7 的(1)可知，存在 $(y^*,z^*)\in\mathscr{Y}^*\times\mathscr{Z}^*$ 和 $(y^*,z^*)\neq\boldsymbol{0}$，使得

$$\langle y^*,y'\rangle+\langle z^*,z'+q_1-z\rangle\geqslant\langle y^*,\bar{y}\rangle+\big\langle z^*,\langle x^*,\bar{x}-x\rangle p_1\big\rangle$$
$$\forall y'\in Y,\ \forall z'\in G(y'),\ \forall q_1\in K_1,\ \forall\bar{x}\in S,\ \forall\bar{y}\in F(\bar{x}). \qquad (6.4.14)$$

在(6.4.14)中令 $z'=z$，$\bar{x}=x$，$y'=y$，$\bar{y}=y$，得到

$$\langle z^*,q_1\rangle\geqslant 0\quad \forall q_1\in K_1.$$

易知有 $z^*\neq\boldsymbol{0}$(因为若 $z^*=\boldsymbol{0}$，则从 $(y^*,z^*)\neq\boldsymbol{0}$ 得 $y^*\neq\boldsymbol{0}$. 由(6.4.14)可知

$$\langle y^*,y'\rangle\geqslant\langle y^*,\bar{y}\rangle\quad \forall y'\in Y,\ \forall\bar{y}\in F(\bar{x}),$$

故得对任意的 $y'\in Y$ 有 $\langle y^*, y'-y\rangle \geqslant 0$. 类似于定理 6.2.4 的证明，导致与 $y\in \text{int}Y$ 相矛盾. 于是，得知有 $z^*\in K_1^*\setminus\{\boldsymbol{0}\}$. 对任意的 $y'\in Y$ 和任意的 $z'\in G(y')$ 在(6.4.14) 中令 $\bar{y}=y$, $\bar{x}=x$, $q_1=\boldsymbol{0}$，则有

$$\langle z^*, z'-z\rangle \geqslant \langle -y^*, y'-y\rangle \quad \forall y'\in Y, \forall z'\in G(y').$$

令 $\bar{y}^* = -y^*/\langle q_1^*, p_1\rangle$, $q_1^* = z^*$，注意 $\langle q_1^*, p_1\rangle > 0$，上式即为

$$\langle q_1^*, z'-z\rangle \geqslant \langle \bar{y}^*, y'-y\rangle\langle q_1^*, p_1\rangle$$
$$\forall y'\in Y, \forall z'\in G(y').$$

由此，根据定理 6.2.11 的(2)得 $\bar{y}^*\in \partial G(y,z)_{p_1}$. 对任意的 $\bar{x}\in S$ 和任意的 $\bar{y}\in F(\bar{x})$ 在(6.3.14) 中令 $q_1=\boldsymbol{0}$, $y'=y$, $z'=z$，可得

$$\langle \bar{y}^*, \bar{y}-y\rangle \geqslant \langle x^*, \bar{x}-x\rangle \quad \forall \bar{x}\in S, \forall \bar{y}\in F(\bar{x}).$$

于是，从定理 6.2.11 的(2)便得 $x^*\in \partial\theta\big(x, \langle \bar{y}^*, y\rangle\big)_1$，由 $\bar{y}^*$ 的任意性，故知(6.4.9) 成立. □

若 $\mathscr{X}=R$, $\mathscr{Y}=R^m$, $F: S\to R^m$ 是单值可微向量函数，则(6.4.9) 成为

$$\partial_w G\circ F(x,z)_{p_1} = \left\{y^*\cdot JF(x) \mid y^*\in \partial_w G(y,z)_{p_1}\right\}.$$

定义 6.4.2　设 $Y\subset\mathscr{Y}$ 是 K-凸集，$G: Y\to 2^{\mathscr{Z}}$, $y\mapsto G(y)$ 是集值映射. 若对任意的 $y^1, y^2\in Y$，任意的 $\lambda\in(0,1)$，存在 $y_\lambda\in Y$, $k\in K_1$，使得 $\lambda y^1+(1-\lambda)y^2 = y_\lambda + k$ 有

$$\lambda G(y^1)+(1-\lambda)G(y^2)\subset G(y_\lambda)+K_1,$$

则称集值映射 G 在 Y 上关于 K 是 K_1-凸的.

显然，若 Y 是凸集，G 在 Y 上关于 K 是 K_1-凸的，则 G 在 Y 上也是 K_1-凸的. 反过来，若 Y 是凸集，G 在 Y 上是 K_1-凸的，则 G 在 Y 上关于 K 也是 K_1-凸的.

定理 6.4.5　设 $S\subset\mathscr{X}$ 是非空凸集，集合 $F(x)\subset\mathscr{Y}$ $(\forall x\in S)$ 是非空的，$Y=F(S)=\bigcup\limits_{x\in S}F(x)$，集合 $G(y)\subset\mathscr{Z}$ $(\forall y\in Y)$ 是

非空的，$x \in S$，$y \in F(x)$，$z \in G(y)$，$y \in \mathrm{int}Y$，$p_1 \in \mathrm{int}K_1$. 又设集值映射 $F: S \to 2^{\mathscr{Y}}$，$x \mapsto F(x)$ 在 S 上是 K-凸的，集值映射 $G: Y \to 2^{\mathscr{Z}}$，$y \mapsto G(y)$ 在 Y 上关于 K 是 K_1-凸的. 对于集值映射

$$G \circ F: S \to 2^{\mathscr{Z}},\ x \mapsto G \circ F(x) = G(F(x)) = \bigcup_{y \in F(x)} G(y),$$

若 G 在 y 处是 K-有界逆开的或连通的，则

$$\partial_w G \circ F(x, z)_{p_1} = \{\partial\theta_1(x, \langle y^*, y\rangle)_1 \mid y^* \in \partial_w G_1(y, z)_{p_1}\},$$

其中集值函数 $\theta_1: S \to 2^R$，$x \mapsto \theta_1(x) = \{\langle y^*, y\rangle \mid y \in F(x) + K\}$，集值映射 $G_1: Y + K \to 2^{\mathscr{Y}}$，$y \mapsto G_1(y)$. 若存在 $y_1 \in Y$，$q_1 \in K$ 使得 $y = y_1 + q_1$，则 $G_1(y) = G(y_1)$.

证明 作集值映射 $F_1: S \to 2^{\mathscr{Y}}$，$x \mapsto F_1(x)$，$F_1(x) = F(x) + K$. 由于 F 在 S 上是 K-凸的，则 F_1 在 S 上是凸的. 由 G_1 的定义可得：

$$\begin{aligned} G \circ F(x) &= \bigcup_{y \in F(x)} G(y) = \bigcup_{y \in F(x)+K} G_1(y) \\ &= \bigcup_{y \in F_1(x)} G_1(y) = G_1 \circ F_1(x). \end{aligned}$$

因为 G 在 Y 上关于 K 是 K_1-凸的，按定义 6.4.2 可以推得 G_1 在 $Y + K$ 上是 K_1-凸的，且由 F 在 S 上是 K_1-凸的知 $Y + K$ 是凸集. 由 $y \in \mathrm{int}Y$ 有 $y \in \mathrm{int}(Y + K)$. 显然，$G_1$ 在 y 处是 K_1-有界逆开的或连通的. 由定理 6.4.4 可知结论成立. □

§6.5 锥方向导数和锥弱方向导数

我们知道，设 $S \subset \mathscr{X}$ 是非空凸集，集值映射 $\psi: S \to 2^{\mathscr{Y}}$，$x \mapsto \psi(x)$ 在 S 上是 K-凸的，则 K-epiψ 是凸集. 又设点 $x^0 \in S$，$y^0 \in \psi(x^0)$，α 是实数，则

$$T_{K\text{-epi}\psi}(x^0, y^0) = \mathrm{cl}\left\{\bigcup_{\alpha>0} \alpha[K\text{-epi}\psi - (x^0, y^0)]\right\} \quad (6.5.1)$$

是 K-epiψ 在点 $(\boldsymbol{x}^0, \boldsymbol{y}^0)$ 处的切锥. (6.5.1) 等价于: $(\boldsymbol{d}, \boldsymbol{r}) \in T_{K\text{-epi}\psi}(\boldsymbol{x}^0, \boldsymbol{y}^0)$ 当且仅当存在序列 $\{\boldsymbol{d}^k\} \subset \mathscr{X}$, $\{\boldsymbol{r}^k\} \subset \mathscr{Y}$ 和 $\{\lambda_k\} \subset (0, +\infty)$, 当 $k \to \infty$ 时有 $\boldsymbol{d}^k \to \boldsymbol{d}$, $\boldsymbol{r}^k \to \boldsymbol{r}$, $\lambda_k \to 0$ 使得 $\boldsymbol{y}^0 + \lambda_k \boldsymbol{r}^k \in \psi(\boldsymbol{x}^0 + \lambda_k \boldsymbol{d}^k) + K$ $(k = 1, 2, \cdots)$. 由此, 从 (6.5.1) 得知

$$K\text{-epi}\psi - (\boldsymbol{x}^0, \boldsymbol{y}^0) \subset T_{K\text{-epi}\psi}(\boldsymbol{x}^0, \boldsymbol{y}^0).$$

定义 6.5.1 设集合 $S \subset \mathscr{X}$ 非空, $\psi: S \to 2^{\mathscr{Y}}$ 是集值映射, 点 $\boldsymbol{x}^0 \in S$, $\boldsymbol{y}^0 \in \psi(\boldsymbol{x}^0)$, $\boldsymbol{d} \in \mathscr{X}$. 记集合

$$\Psi(\boldsymbol{x}^0, \boldsymbol{y}^0, \boldsymbol{d}) = \{\boldsymbol{r} \in \mathscr{Y} \mid (\boldsymbol{d}, \boldsymbol{r}) \in T_{K\text{-epi}\psi}(\boldsymbol{x}^0, \boldsymbol{y}^0)\}. \tag{6.5.2}$$

(1) 若 $\boldsymbol{y}'(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d}) \in \Psi(\boldsymbol{x}^0, \boldsymbol{y}^0, \boldsymbol{d})$, 并且不存在 $\boldsymbol{y} \in \Psi(\boldsymbol{x}^0, \boldsymbol{y}^0, \boldsymbol{d})$ 使得

$$\boldsymbol{y}'(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d}) - \boldsymbol{y} \in K \setminus \{\boldsymbol{0}\},$$

则称 $\boldsymbol{y}'(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d})$ 是集值映射 ψ 在点 $(\boldsymbol{x}^0, \boldsymbol{y}^0)$ 处沿方向 $\boldsymbol{d}$ 的 K-方向导数. ψ 在 $(\boldsymbol{x}^0, \boldsymbol{y}^0)$ 处沿 $\boldsymbol{d}$ 的所有 K-方向导数组成的集合记作 $\psi'(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d})$.

(2) 若 $\boldsymbol{y}'(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d}) \in \Psi(\boldsymbol{x}^0, \boldsymbol{y}^0, \boldsymbol{d})$, 并且不存在 $\boldsymbol{y} \in \Psi(\boldsymbol{x}^0, \boldsymbol{y}^0, \boldsymbol{d})$ 使得

$$\boldsymbol{y}'(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d}) - \boldsymbol{y} \in \text{int}K,$$

则称 $\boldsymbol{y}'(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d})$ 是集值映射 ψ 在点 $(\boldsymbol{x}^0, \boldsymbol{y}^0)$ 处沿方向 $\boldsymbol{d}$ 的 K-弱方向导数. ψ 在 $(\boldsymbol{x}^0, \boldsymbol{y}^0)$ 处沿方向 $\boldsymbol{d}$ 的所有 K-弱方向导数组成的集合记作 $\psi'_w(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d})$.

显然有 $\psi'(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d}) \subset \psi'_w(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d})$. 当 $\mathscr{Y} = R$, $K = R_+$ 时, 则 $\psi'(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d}) = \psi'_w(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d})$.

注 6.5.1 特别地, 若 $\mathscr{X} = \mathscr{B}$ 是 Banach 空间, $\psi = f: S \to R \cup \{+\infty\}$ 是实值凸函数, 并且 f 在 $\boldsymbol{x}^0 \in S$ 处满足 Lipschitz 条件, 则由定义 5.2.2、定义 5.3.1 和定义 5.4.1 可知, 这里的锥方向导数 $f'(\boldsymbol{x}^0, f(\boldsymbol{x}^0); \boldsymbol{d})$ 与 D- 切导数或 C- 切导数是一致的. 并且, 由

定理 5.4.2 还可知，锥方向导数 $f'(\boldsymbol{x}^0, f(\boldsymbol{x}^0); \boldsymbol{d})$ 等于第 4 章定义的 G- 方向导数，即 $f'(\boldsymbol{x}^0, f(\boldsymbol{x}^0); \boldsymbol{d}) = f^0(\boldsymbol{x}^0; \boldsymbol{d})$.

以下讨论锥方向导数和锥弱方向导数与锥次微分和锥弱次微分的关系.

定义 6.5.2 设 $Y \subset \mathscr{Y}$ 是非空集合，$K \subset \mathscr{Y}$ 是闭凸锥.

(1) 若存在点 $\boldsymbol{a} \in \mathscr{Y}$ 使得 $Y \subset \boldsymbol{a} + K$，则称 Y 是K- 有界集，或 Y 是K- 有界的.

(2) 若 $Y + K$ 是闭集，则称 Y 是K- 闭集，或 Y 是K- 闭的.

引理 6.5.1 设集合 $Y \subset \mathscr{Y}$ 非空. 若 Y 是 K- 闭的和 K- 有界的，则 Y 是 K- 半紧的.

证明 见文献[27]的命题 3.3.5. ▯

引理 6.5.2 设集合 $Y \subset \mathscr{Y}$ 非空，$K \subset \mathscr{Y}$ 是内部非空的尖闭凸锥，Y 是 K- 半紧的.

(1) $Y \subset \mathscr{E}(Y, K) + K$.

(2) $Y \subset \mathscr{E}_w(Y, K) + \operatorname{int}K \cup \{\boldsymbol{0}\}$.

证明 见文献[27]的定义 3.5.1 和定理 3.5.4. ▯

设 $\boldsymbol{x} \in S$, $\boldsymbol{y} \in \psi(\boldsymbol{x})$，记

$$\widehat{D} = \{\boldsymbol{d} \in \mathscr{X} \mid \Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}) \neq \varnothing\}.$$

若 $S = \mathscr{X}$，则显然有 $\widehat{D} = \mathscr{X}$.

定理 6.5.3 设 $S \subset \mathscr{X}$ 是非空凸集，$\psi: S \to 2^{\mathscr{Y}}$ 是集值映射，$\boldsymbol{x} \in S$, $\boldsymbol{y} \in \psi(\boldsymbol{x})$, $\widehat{D} = S$, $\boldsymbol{p} \in \operatorname{int}K$. 若 ψ 在 S 上是 K- 凸的，并且 $\Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0})$ 是K-有界的，则对任意的 $\boldsymbol{d} \in \mathscr{X}$ 有 $\psi'(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}) \neq \varnothing$ 和 $\psi'_w(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}) \neq \varnothing$，并且

$$\begin{aligned}
&\partial\psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}} \\
&\supset \left\{\boldsymbol{x}^* \in \mathscr{X}^* \;\middle|\; \boldsymbol{0} \in \mathscr{E}\left(\bigcup_{\boldsymbol{d} \in \mathscr{X}} [\psi'(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}) - \langle \boldsymbol{x}^*, \boldsymbol{d}\rangle \boldsymbol{p}], K\right)\right\}, \\
&\partial_w\psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}} \\
&= \left\{\boldsymbol{x}^* \in \mathscr{X}^* \;\middle|\; \boldsymbol{0} \in \mathscr{E}_w\left(\bigcup_{\boldsymbol{d} \in \mathscr{X}} [\psi'_w(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}) - \langle \boldsymbol{x}^*, \boldsymbol{d}\rangle \boldsymbol{p}], K\right)\right\}.
\end{aligned}$$

证明　取任意的 $d^1, d^2\in\mathscr{X}$,任意的 $r^1, r^2\in\Psi(x, y, d)$ 由(6.5.2)知 (d^1, r^1), $(d^2, r^2)\in T_{K\text{-epi}\phi}(x, y)$. 因为 $T_{K\text{-epi}\phi}(x, y)$ 是凸锥,故 $(d^1+d^2, r^1+r^2)\in T_{K\text{-epi}\phi}(x, y)$,再由(6.5.2)有 $r^1+r^2\in\Psi(x, y, d^1+d^2)$. 因此,得到

$$\Psi(x, y, d^1)+\Psi(x, y, d^2)\subset\Psi(x, y, d^1+d^2)\quad\forall d^1, d^2\in\mathscr{X}.$$

对任意的 $d\in\mathscr{X}$,令 $d^1=d$, $d^2=-d$,取点 $r\in\Psi(x, y, d^2)$,我们有

$$\Psi(x, y, d)+r\subset\Psi(x, y, 0).$$

因为 $\Psi(x, y, 0)$是 K-有界的,所以由上式可知 $\Psi(x, y, d)$也是 K-有界的. 另外,由 $T_{K\text{-epi}\phi}(x, y, d)$是闭的,易知 $\Psi(x, y, d)$也是闭的,又因 K 是闭的,从而 $\Psi(x, y, d)$是 K-闭的. 由此,根据引理 6.5.1 可知 $\Psi(x, y, d)$是 K-半紧的,再由引理 6.2.5 得知 $\psi'(x, y; d)\neq\varnothing$ 和 $\psi'_w(x, y; d)\neq\varnothing$. 于是,根据引理 6.5.2 得到

$$\Psi(x, y, d)\subset\psi'(x, y; d)+K, \tag{6.5.3}$$

$$\Psi(x, y, d)\subset\psi'_w(x, y; d)+K. \tag{6.5.4}$$

记集合

$$S_1=\left\{x^*\in\mathscr{X}^*\ \middle|\ 0\in\mathscr{E}\left(\bigcup_{(d, r)\in T_{K\text{-epi}\phi}(x, y)}\left[r-\langle x^*, d\rangle p\right], K\right)\right\}, \tag{6.5.5}$$

$$S_2=\left\{x^*\in\mathscr{X}^*\ \middle|\ 0\in\mathscr{E}\left(\bigcup_{d\in\mathscr{X}}\left[\psi'(x, y; d)-\langle x^*, d\rangle p\right], K\right)\right\}, \tag{6.5.6}$$

$$S_3=\left\{x^*\in\mathscr{X}^*\ \middle|\ 0\in\mathscr{E}_w\left(\bigcup_{(d, r)\in T_{K\text{-epi}\phi}(x, y)}\left[r-\langle x^*, d\rangle p\right], K\right)\right\}, \tag{6.5.7}$$

$$S_4=\left\{x^*\in\mathscr{X}^*\ \middle|\ 0\in\mathscr{E}_w\left(\bigcup_{d\in\mathscr{X}}\left[(\psi'_w(x, y; d)-\langle x^*, d\rangle p\right], K\right)\right\}. \tag{6.5.8}$$

先证明 $S_1 = S_2$. 设 $x^* \in S_2$,假若 $x^* \notin S_1$,则存在 $(d', r') \in T_{K\text{-epi}\psi}(x, y)$,使得

$$\theta - \left(r' - \langle x^*, d' \rangle p\right) \in K \setminus \{\theta\}. \qquad (6.5.9)$$

另外,由(6.5.2)和(6.5.3)知 $r' \in \psi'(x, y; d') + K$,故存在 $q' \in K$, $y' \in \psi'(x, y; d')$ 使 $r' = y' + q'$. 再由(6.5.9)有

$$\theta - (y' - \langle x^*, d' \rangle p) \in K \setminus \{\theta\} + q' \subset K \setminus \{\theta\},$$

于是得知 $x^* \notin S_2$,这与已设矛盾. 反之,设 $x^* \in S_1$,则不存在 $(d, r) \in T_{K\text{-epi}\psi}(x, y)$ 使得

$$-\left(r - \langle x^*, d \rangle p\right) \in K \setminus \{\theta\}. \qquad (6.5.10)$$

由定义 6.5.1 知 $\psi'(x, y; d) \subset \Psi(x, y, d)$,故再由(6.5.10)可知不存在 $r \in \psi'(x, y; d)$,使 $-\left(r - \langle x^*, d \rangle p\right) \in K \setminus \{\theta\}$,从而 $x^* \in S_2$.

同证明 $S_1 = S_2$ 类似,可由(6.5.4)推得 $S_3 = S_4$. 按定义 6.1.2 可知 $S_1 \subset S_3$,从而有 $S_2 \subset S_4$. 现证 $S_2 \subset \partial\psi(x, y)_p$. 设 $x^* \in S_2 = S_1$,从(6.5.1)知对任意的 $x' \in S$ 和 $y' \in \psi(x')$ 有 $(x' - x, y' - y) \in T_{K\text{-epi}\psi}(x, y)$. 于是,由(6.5.5)可得 $y - y' + \langle x^*, x' - x \rangle p \notin K \setminus \{\theta\}$,再由定义 6.1.3 的(1)便知 $x^* \in \partial\psi(x, y)_p$. 类似地,由(6.5.7)和(6.5.8)可以推得 $S_4 \subset \partial_w\psi(x, y)_p$.

最后,证明 $\partial_w\psi(x, y)_p \subset S_4$. 用反证法,假设 $x^* \in \partial\psi(x, y)_p$ 而 $x^* \notin S_4$,则有 $x^* \notin S_3$. 由此,存在 $(d', r') \in T_{K\text{-epi}\psi}(x, y)$ 使得

$$\theta - \left(r' - \langle x^*, d' \rangle p\right) \in \mathrm{int}K, \qquad (6.5.11)$$

因而有序列 $\{d^k\} \subset \mathscr{X}$, $\{r^k\} \subset \mathscr{Y}$ 和 $\{\lambda_k\} \subset (0, +\infty)$,当 $k \to \infty$ 时,有 $d^k \to d'$, $r^k \to r'$, $\lambda_k \to 0$,使得

$$y + \lambda_k r^k \in \psi(x + \lambda_k d^k) + K, \quad k = 1, 2, \cdots.$$

据此:存在 $y_k \in \psi(x + \lambda d^k)$ 和 $p^k \in K$ 使 $y_k + p^k = y + \lambda_k r^k$ $(k = 1, 2, \cdots)$,或有

$$r^k = (y_k - y + p^k)/\lambda_k, \quad k = 1, 2, \cdots.$$

令 $x^k = x + \lambda_k d^k$，则 $d^k = (x^k - x)/\lambda_k \quad (k = 1, 2, \cdots)$，$y^k \in \psi(x^k)$. 因为当 $k \to \infty$ 时，$r^k \to r'$，$d^k \to d'$，故对充分大的 k'，由(6.5.11)得知

$$-(y^{k'} - y + p^{k'})/\lambda_{k'} + \langle x^*, (x^{k'} - x)/\lambda_{k'} \rangle p \in \operatorname{int}K,$$

从而有 $y - y^{k'} + \langle x^*, x^{k'} - x \rangle p \in \operatorname{int}K$. 于是，根据定义 6.1.3 的(2)得 $x^* \notin \partial\psi(x, y)_p$，导致与已设矛盾. □

定理 6.5.4　设 $S \subset \mathscr{X}$ 是非空凸集，$\psi: S \to 2^{\mathscr{Y}}$ 是集值映射，$x \in S$，$y \in \psi(x)$，$\hat{D} = S$，$p \in \operatorname{int}K$. 若 ψ 在 S 上是 K-凸的，并且 $\Psi(x, y, 0)$ 是 K-有界的，则对任意的 $d \in \mathscr{X}$，有 $\psi'(x, y; d)$ 和 $\psi'_w(x, y; d)$ 关于 d 在 $\mathscr{X}$ 上是 K-凸的.

证明　由定理 6.5.3 的证明可得

$$\Psi(x, y, d^1) + \Psi(x, y, d^2) \subset \Psi(x, y, d^1 + d^2)$$

$$\forall d^1, d^2 \in \mathscr{X}.$$

再由(6.5.3)和(6.5.4)得到

$$\psi'(x, y; d^1) + \psi'(x, y; d^2) \subset \psi'(x, y; d^1 + d^2) + K,$$

$$\psi'_w(x, y; d^1) + \psi'_w(x, y; d^2) \subset \psi'_w(x, y; d^1 + d^2) + K.$$

因此，定理的结论成立. □

定理 6.5.5　设 $S \subset \mathscr{X}$ 是非空凸集，$\psi: S \to 2^{\mathscr{Y}}$ 是集值映射，$x \in S$，$y \in \psi(x)$，$\hat{D} = S$，$p \in \operatorname{int}K$. 又设 ψ 在 S 上是 K-凸的，并且 $\Psi(x, y, 0)$ 是 K-有界的.

(1) 若存在 $q^* \in \operatorname{int}K^*$，使得

$$\langle q^*, r \rangle \geqslant \langle x^*, d \rangle \langle q^*, p \rangle \quad \forall (d, r) \in T_{K\text{-epi}\psi}(x, y),$$

则 $x^* \in \partial\psi(x, y)_p$.

(2) 若存在 $q^* \in \operatorname{int}K^*$，使得

$$\langle q^*, y'\rangle \geqslant \langle x^*, d\rangle\langle q^*, p\rangle$$

$$\forall d \in \mathscr{X}, \forall y' \in \psi'(x, y; d),$$

则 $x^* \in \partial\psi(x, y)_p$.

(3) 若存在 $q^* \in K^*\setminus\{0\}$,使得

$$\langle q^*, r\rangle \geqslant \langle x^*, d\rangle\langle q^*, p\rangle \quad \forall (d, r) \in T_{K\text{-epi}\psi}(x, y),$$

则 $x^* \in \partial_w\psi(x, y)_p$.

(4) 若存在 $q^* \in K^*\setminus\{0\}$,使得

$$\langle q^*, y'\rangle \geqslant \langle x^*, d\rangle\langle q^*, p\rangle$$

$$\forall d \in \mathscr{X}, \forall y' \in \psi'_w(x, y; d),$$

则 $x^* \in \partial_w\psi(x, y)_p$.

证明 (1) 因为对任意的 $x' \in S$, $y' \in \psi(x')$,有 $(x' - x, y' - y) \in T_{K\text{-epi}\psi}(x, y)$,所以根据定理 6.2.11 的(1)可得结论.

(2) 反证:假若 $x^* \notin \partial\psi(x, y)_p$,由定理 6.5.3 可知,存在 $\bar{y} \in \psi'(x, y; d)$ 使得

$$0 - \big(\bar{y} - \langle x^*, d\rangle p\big) \in K\setminus\{0\}.$$

再由 $q^* \in \text{int}K^*$,得 $\langle q^*, \bar{y}\rangle < \langle x^*, d\rangle\langle q^*, p\rangle$,与已知矛盾.

(3) 与(1)的证明类似.

(4) 与(2)的证明类似. □

定理 6.5.6 设 $S \subset \mathscr{X}$ 是非空凸集,$\psi: S \to 2^{\mathscr{Y}}$ 是集值映射,$x \in S$, $y \in \psi(x)$, $\hat{D} = S$, $p \in \text{int}K$. 又设 ψ 在 S 上是 K-凸的,并且 $\Psi(x, y, 0)$ 是 K-有界的. 若 $x^* \in \partial_w\psi(x, y)_p$,则存在 $q^* \in K^*\setminus\{0\}$,使得

$$\langle q^*, y'\rangle \geqslant \langle x^*, d\rangle\langle q^*, p\rangle$$

$$\forall d \in \mathscr{X}, \forall y' \in \psi'(x, y; d),$$

$$\langle q^*, y'\rangle \geqslant \langle x^*, d\rangle\langle q^*, p\rangle$$

$$\forall d \in \mathscr{X},\ \forall y' \in \psi'_w(x,\ y;\ d),$$

$$\langle q^*,\ r\rangle \geqslant \langle x^*,\ d\rangle\langle q^*,\ p\rangle \quad \forall (d,\ r) \in T_{K\text{-epi}\psi}(x,\ y).$$

证明 只需证明最后一个式子成立. 设 $x^* \in \partial_w\psi(x,\ y)_p$, 由定理 6.5.4 的证明可知 $x^* \in S_3$. 作集合

$$E = \{y' \in \mathscr{Y} \mid y' = -r + \langle x^*,\ d\rangle,\ (d,\ r) \in T_{K\text{-epi}\psi}(x,\ y)\},$$

则 E 是凸集, 并且 $E \cap \text{int}K = \varnothing$. 由此, 利用定理 1.3.7 的(1) 可知, 存在 $q^* \neq 0$ 使得

$$\langle q^*,\ q\rangle \geqslant \langle q^*,\ y'\rangle \quad \forall q \in K,\ \forall y' \in E.$$

由于 $0 \in E \cap K$, 故对任意的 $q \in K$, 有 $\langle q^*,\ q\rangle \geqslant 0$, 因此 $q^* \in K^* \setminus \{0\}$. 再由 $0 \geqslant \langle q^*,\ y'\rangle (y' \in E)$, 即得结论. ▯

定理 6.5.7 设 $S \subset \mathscr{X}$ 是非空凸集, $\psi: S \to 2^{\mathscr{Y}}$ 是集值映射, $x \in S$, $y \in \psi(x)$, $\hat{D} = S$, $p \in \text{int}K$. 又设 ψ 在 S 上是 K-凸的, 并且 $\Psi(x,\ y,\ 0)$ 是 K-有界的. 若存在 $\bar{x}^* \in \partial_w\psi(x,\ y)_p$ 和 $\bar{y} \in \psi'(x,\ y;\ \bar{d})$, 使得 $\bar{y} = \langle \bar{x}^*,\ \bar{d}\rangle p$, 则

$$\bar{y} = \left(\max\{\langle x^*,\ \bar{d}\rangle \mid x^* \in \partial_w\psi(x,\ y)_p\}\right)p. \quad (6.5.12)$$

证明 作集合

$$E = \{y' \in \mathscr{Y} \mid y' = -\bar{y} + \langle x^*,\ \bar{d}\rangle p,\ x^* \in \partial_w\psi(x,\ y)_p\},$$

注意: $\bar{y}$ 和 $\bar{d}$ 是给定的. 从定理 6.3.2 知 $\partial_w\psi(x,\ y)_p$ 是凸的, 故 E 是凸的. 由定理 6.5.3 可推得 $E \cap \text{int}K = \varnothing$, 再利用定理 1.3.7, 即知存在 $q^* \neq 0$ 使得

$$\langle q^*,\ q\rangle \geqslant \langle q^*,\ y'\rangle \quad \forall q \in K,\ \forall y' \in E.$$

由于 $0 \in E \cap K$, 故对任何 $q \in K$ 有 $\langle q^*, q\rangle \geqslant 0$, 因此 $q^* \in K^* \setminus \{0\}$. 上式对任何 $y' = -\bar{y} + \langle x^*,\ \bar{d}\rangle p \in E$, 有

$$\langle q^*,\ \bar{y}\rangle \geqslant \langle x^*,\ \bar{d}\rangle\langle q^*,\ p\rangle \quad \forall x^* \in \partial_w\psi(x,\ y)_p.$$

从已知 $\bar{y}=\langle\bar{x}^*, \bar{d}\rangle p$ 代入上式,得

$$\langle\bar{x}^*, \bar{d}\rangle \geqslant \langle x^*, \bar{d}\rangle \quad \forall x^* \in \partial_w\psi(x, y)_p,$$

因而(6.5.12)成立.▯

定义 6.5.3 设 $S\subset\mathscr{X}$ 是非空凸集,ψ: $S\to 2^{\mathscr{Y}}$ 是集值映射,点 $x^0\in S$, $y^0\in\psi(x^0)$. 若 ψ 在 S 上是 K-凸的,则称集合

$$\begin{aligned} &N_{K\text{-epi}\psi}(x^0, y^0) \\ &= \{(x^*, y^*)\in\mathscr{X}^*\times\mathscr{Y}^* \mid \langle(x^*, y^*), (d, r)\rangle\leqslant 0, \\ &\qquad\qquad \forall(d, r)\in T_{K\text{-epi}\psi}(x^0, y^0)\} \end{aligned}$$

是集值映射 ψ 在点 (x^0, y^0) 处的法锥.

定理 6.5.8 设 $S\subset\mathscr{X}$ 是非空凸集,ψ: $S\to 2^{\mathscr{Y}}$ 是集值映射,$x\in S$, $y\in\psi(x)$, $\hat{D}=S$, $p\in\text{int}K$. 又设 ψ 在 S 上是 K-凸的,并且 $\Psi(x, y, 0)$ 是 K-有界的.

(1) $x^*\in\partial_w\psi(x, y)_p$ 当且仅当存在 $q^*\in K^*\setminus\{0\}$,使得

$$(\langle q^*, p\rangle x^*, -q^*)\in N_{K\text{-epi}\psi}(x, y).$$

(2) 若存在 $q^*\in\text{int}K^*$,使得

$$(\langle q^*, p\rangle x^*, -q^*)\in N_{K\text{-epi}\psi}(x, y),$$

则 $x^*\in\partial\psi(x, y)_p$.

证明 由定义 6.5.3、定理 6.5.5 和定理 6.5.6 可得证.▯

第 7 章　一般集值映射的锥弱次微分

在上一章，我们研究了锥凸集值映射的锥次微分问题. 本章将对一般的集值映射引进两种锥弱次微分概念，并讨论它们的一些相应的性质.

§7.1　有效 Hahn-Banach 定理

在第 4 章，我们看到了 Hahn-Banach 定理在证明 Lipschitz 函数的次微分的存在性以及其他性质中起到了关键的作用. 现在先证明关于集值映射的有效 Hahn-Banach 定理，为后面研究一般集值映射的锥弱次微分打下基础.

设 $\mathscr{X}$ 和 $\mathscr{Y}$ 是局部凸线性拓扑空间，$\mathscr{X}^*$ 和 $\mathscr{Y}^*$ 分别是 $\mathscr{X}$ 和 $\mathscr{Y}$ 的对偶空间，集合 $S \in \mathscr{X}$，$\psi: S \to 2^Y$，$\boldsymbol{x} \mapsto \psi(\boldsymbol{x})$ 是集值映射. 又设 $K \subset \mathscr{Y}$ 是内部非空的尖闭凸锥，$\boldsymbol{p} \in \mathrm{int}K$，$K^*$ 是 K 的对偶锥.

定义 7.1.1　设集合 $S \subset \mathscr{X}$ 非空，点 $\boldsymbol{x}^0 \in \mathrm{int}S$. 若存在一点 $\boldsymbol{a}$ 和 $\boldsymbol{x}^0$ 的一个邻域 $U(\boldsymbol{x}^0)$，使得 $\psi(\boldsymbol{x}) \subset \boldsymbol{a} + K\ \left(\forall \boldsymbol{x} \in U(\boldsymbol{x}^0)\right)$，则称集值映射 ψ 在点 $\boldsymbol{x}^0$ 附近是*局部 K-有界的*.

引理 7.1.1　设集合 $S \subset \mathscr{X}$ 非空，$\boldsymbol{x}^0 \in \mathrm{int}S$. 若 ψ 在 $\boldsymbol{x}^0$ 附近是局部$(-K)$-有界的，则 $\mathrm{int}(K\text{-epi}\psi) \neq \varnothing$，其中 K-epiψ 是 ψ 在 S 上的 K-上图象.

证明　与引理 6.2.1 的证明类似. □

定理 7.1.2　设 $\mathscr{L} \subset \mathscr{X}$ 是非空线性子空间，ψ 在 $\mathscr{X}$ 上是

K-凸的，ψ 在 $x^0 \in \mathscr{X}$ 附近是局部$(-K)$-有界的(或 K-有界逆开，或连通的). 若存在点 $x_0^* \in \mathscr{X}^*$，使得

$$0 \in \mathscr{E}_w\Big(\bigcup_{x\in\mathscr{L}}\big[\psi(x) - \langle x_0^*, x\rangle p\big], K\Big),$$

则存在线性泛函 $\bar{x}^* \in \mathscr{X}^*$ 使得

$$\langle x_0^*, x\rangle = \langle \bar{x}^*, x\rangle \quad \forall x \in \mathscr{L},$$

并且

$$0 \in \mathscr{E}_w\Big(\bigcup_{x\in\mathscr{X}}\big[\psi(x) - \langle \bar{x}^*, x\rangle p\big], K\Big).$$

证明 设集合

$$S_1 = \{(x, y) \mid x \in \mathscr{X}, y \in \psi(x) + K\},$$

$$S_2 = \{(x, y) \mid x \in \mathscr{L}, y = \langle x_0^*, x\rangle p\},$$

则 S_1 和 S_2 是 $\mathscr{X}\times\mathscr{Y}$ 中的子集，并且易知 $S_2 \neq \varnothing$. 另外，由引理 7.2.1(或引理 6.2.1 或引理 6.2.3)可得 $\mathrm{int}S_1 \neq \varnothing$. 现在证明

$$S_2 \cap \mathrm{int}S_1 = \varnothing. \tag{7.1.1}$$

事实上，假若有 $(x', y') \in S_2 \cap \mathrm{int}S_1$，因为 $p \neq 0$，则存在 $\lambda > 0$ 使得 $(x', y' - \lambda p) \in S_1$. 因此，有 $x' \in \mathscr{L}$，$y'' \in \psi(x')$ 和 $q' \in K$，使得

$$\langle x_0^*, x'\rangle p - \lambda p = y'' + q'.$$

由于 $p \in \mathrm{int}K$，故有

$$0 - \big(y'' - \langle x_0^*, x'\rangle p\big) = \lambda p + q' \in \mathrm{int}K.$$

据此，由定义 6.1.2 的(2)得

$$0 \notin \mathscr{E}_w\Big(\bigcup_{x\in\mathscr{L}}\big[\psi(x) - \langle x_0^*, x\rangle p\big], K\Big),$$

这导致与已知矛盾. 故(7.1.1)成立. 现由(7.1.1)，利用定理 1.3.7 得知，存在 $(x^*, y^*) \in \mathscr{X}^* \times \mathscr{Y}^*$ 和 $(x^*, y^*) \neq 0$使得

$$\langle x^*, x\rangle + \langle y^*, y\rangle + q \geqslant \langle x^*, x_v\rangle + \langle y^*, \langle x_0^*, x_v\rangle p\rangle$$

$$\forall x \in \mathscr{X}, \forall y \in \psi(x), \forall q \in K, \forall x_v \in \mathscr{L}. \quad (7.1.2)$$

由于已知 $\boldsymbol{0} \in \bigcup_{x\in\mathscr{L}} [\psi(x) - \langle x_0^*, x\rangle p]$，从而存在 $\bar{x} \in \mathscr{L}$ 使得 $\bar{y} \in \psi(\bar{x})$ 和 $\bar{y} = \langle x_0^*, \bar{x}\rangle p$，因此 $(\bar{x}, \bar{y}) \in S_2 \cap S_1$. 由(7.1.2)可得

$$\langle y^*, q\rangle \geqslant 0 \quad \forall q \in K.$$

因为 $\mathscr{L}$ 是线性子空间，$y^* \neq \boldsymbol{0}$，所以 $y^* \in K^* \backslash \{\boldsymbol{0}\}$. 再由 $p \in \text{int}K$ 和定理 1.4.11 的(1)可得 $\langle y^*, p\rangle > 0$. 又由于 $\mathscr{L}$ 是线性子空间，故由(7.1.2)得到

$$\langle x^*, x_v\rangle + \langle x_0^*, x_v\rangle\langle y^*, p\rangle = 0 \quad \forall x_v \in \mathscr{L}.$$

设 $\bar{x}^* = -x^*/\langle y^*, p\rangle$，从上式有

$$\langle \bar{x}^*, x_v\rangle = \langle x_0^*, x_v\rangle \quad \forall x_v \in \mathscr{L},$$

再由(7.1.2)可得

$$\langle y^*, y\rangle - \langle \bar{x}^*, x\rangle\langle y^*, p\rangle \geqslant 0 \quad \forall x \in \mathscr{X}, \forall y \in \psi(x). \tag{7.1.3}$$

现假设

$$\boldsymbol{0} \notin \mathscr{E}_w\Big(\bigcup_{x\in\mathscr{X}} [\psi(x) - \langle \bar{x}^*, x\rangle p], K\Big),$$

根据定义 6.1.2 可知存在 $\bar{x}^0 \in \mathscr{X}$，$\bar{y}^0 \in \psi(\bar{x}^0)$，使得

$$\boldsymbol{0} - \big(\bar{y}^0 - \langle \bar{x}^*, \bar{x}^0\rangle p\big) \in \text{int}K.$$

因为 $y^* \in K^* \backslash \{\boldsymbol{0}\}$，所以得到

$$\langle y^*, -\bar{y}^0 + \langle \bar{x}^*, \bar{x}^0\rangle p\rangle > 0,$$

这与(7.1.3)矛盾. 于是定理得证. ▯

推论 7.1.3　设 $\mathscr{L} \subset \mathscr{X}$ 是非空线性子空间，ψ 在 $\mathscr{X}$ 上是 K-凸的，ψ 在 $x^0 \in \mathscr{X}$ 附近是局部 $(-K)$- 有界的(或 K-有界逆

开,或连通的). 若存在点 $x_0^* \in \mathscr{X}^*$ 和 $q_0^* \in K^* \setminus \{\boldsymbol{0}\}$, 使得

$$\langle q_0^*, y\rangle \geqslant \langle x_0^*, x\rangle\langle q_0^*, p\rangle \quad \forall x \in \mathscr{L}, \forall y \in \psi(x),$$

以及 $\bar{y} = \langle x_0^*, \bar{x}\rangle p$, 其中 $\bar{y} \in \psi(\bar{x})$, 则存在线性泛函 $\bar{x}^* \in \mathscr{X}^*$ 和 $\bar{q}^* \in K^* \setminus \{\boldsymbol{0}\}$, 使得

$$\langle \bar{x}^*, x\rangle = \langle x_0^*, x\rangle \quad \forall x \in \mathscr{L}$$

和

$$\langle \bar{q}^*, y\rangle \geqslant \langle \bar{x}^*, x\rangle\langle \bar{q}^*, p\rangle \quad \forall x \in \mathscr{X}, \forall y \in \psi(x).$$

证明 先证明有 $\boldsymbol{0} \in \mathscr{E}_w\left(\bigcup_{x\in\mathscr{L}}\left[\psi(x) - \langle x_0^*, x\rangle p\right], K\right)$. 因为否则, 由定义 6.1.2 的(2)可知, 存在 $x' \in \mathscr{L}$, $y' \in \psi(x')$, 使得

$$\boldsymbol{0} - \left(y' - \langle x_0^*, x'\rangle p\right) \in \operatorname{int}K.$$

据此, 由 $p \in \operatorname{int}K$ 和 $q_0^* \in K^* \setminus \{\boldsymbol{0}\}$, 我们有

$$\langle q_0^*, -y'\rangle + \langle x_0^*, x'\rangle p > 0,$$

这导致与已知矛盾. 由已证的结果, 根据定理 7.1.2 的证明和 (7.1.3) 可知, 存在线性泛函 $\bar{x}^* \in \mathscr{X}^*$ 和 $\bar{q}^* \in K^* \setminus \{\boldsymbol{0}\}$, 使得本定理的结论成立. ▯

推论 7.1.4 设 $\mathscr{L} \subset \mathscr{X}$ 是非空线性子空间, $f: \mathscr{X} \to R$ 是凸函数, 并且是连续的. 若存在 $x_0^* \in \mathscr{X}^*$ 使得

$$\langle x_0^*, x\rangle \leqslant f(x) \quad \forall x \in \mathscr{L},$$

又 $f(\bar{x}) = \langle x_0^*, \bar{x}\rangle$, 其中 $\bar{x} \in \mathscr{X}$, 则存在线性泛函 $\bar{x}^* \in \mathscr{X}^*$ 使得

$$\langle x_0^*, x\rangle = \langle \bar{x}^*, x\rangle \quad \forall x \in \mathscr{L}$$

和

$$\langle \bar{x}^*, x\rangle \leqslant f(x) \quad \forall x \in \mathscr{X}.$$

证明 由推论 7.1.3 可得证. ▯

为了证明集值映射在一点处的延拓成立, 以下引进集值映射的锥次线性性的概念.

定义 7.1.2 设集合 $S \subset \mathscr{X}$ 非空, $\psi: \mathscr{X} \to 2^{\mathscr{Y}}$ 是集值映射.

(1) 对任意的 $x \in S$ 和 $\alpha > 0$，有 $\psi(\alpha x) = \alpha\psi(x)$（正齐次性）；

(2) 对任意的 $x^1, x^2 \in S$，有 $\psi(x^1) + \psi(x^2) \subset \psi(x^1 + x^2) + K$（次可加性）.

则称集值映射 ψ 在 S 上是 K-次线性的.

显然，ψ 在 S 上是 K-次线性的，则 ψ 在 S 上是 K-凸的.

定理 7.1.5 设集值映射 ψ 在 $\mathscr{X}$ 上是 K-次线性的，ψ 在 $\mathscr{X}$ 中的某一点处附近是局部 $(-K)$-有界的(或 K-有界逆开的，或连通的)，$x^0 \in \mathscr{X}$，$\boldsymbol{0} \in \psi(\boldsymbol{0}) \subset K$. 若有 $q_0^* \in K^* \setminus \{\boldsymbol{0}\}$ 使 $\inf\limits_{y \in \psi(x^0)} \langle q_0^*, y\rangle > -\infty$，则存在线性泛函 $x^* \in \mathscr{X}^*$，使得

$$\inf\langle q_0^*, \psi(x^0)\rangle = \langle x^*, x^0\rangle\langle q_0^*, p\rangle \qquad (7.1.4)$$

和

$$\boldsymbol{0} \in \mathscr{E}_w\Big(\bigcup_{x \in \mathscr{X}}\big[\psi(x) - \langle x^*, x\rangle p\big], K\Big).$$

证明 先考虑 $x^0 \neq \boldsymbol{0}$ 的情况. 从 $p \in \mathrm{int}K$ 和 $q_0^* \in K^* \setminus \{\boldsymbol{0}\}$，由定理 1.4.11 的(1)得 $\langle q_0^*, p\rangle > 0$.

设 $\alpha = \inf\big\langle q_0^*, \psi(x^0)\big\rangle / \langle q_0^*, p\rangle$，则有

$$\langle q_0^*, y\rangle \geqslant \alpha\langle q_0^*, p\rangle \quad \forall y \in \psi(x^0). \qquad (7.1.5)$$

因为 ψ 在 $\mathscr{X}$ 上是 K-次线性的，我们有

$$\psi(x^0) + \psi(-x^0) \subset \psi(\boldsymbol{0}) + K \subset K.$$

由 $q_0^* \in K^* \setminus \{\boldsymbol{0}\}$，则得

$$\langle q_0^*, y\rangle + \langle q_0^*, \bar{y}\rangle \geqslant 0 \quad \forall y \in \psi(x^0), \forall \bar{y} \in \psi(-x^0). \qquad (7.1.6)$$

下面证明有

$$\langle q_0^*, \bar{y}\rangle \geqslant -\alpha\langle q_0^*, p\rangle \quad \forall \bar{y} \in \psi(-x^0). \qquad (7.1.7)$$

事实上，假设(7.1.7)不成立，则存在 $\delta > 0$ 和 $\bar{y}^0 \in \psi(-x^0)$，使得

$$\langle q_0^*, \bar{y}^0\rangle < -\alpha\langle q_0^*, p\rangle - \delta, \qquad (7.1.8)$$

再由 α 的定义可知存在 $y^0 \in \psi(x^0)$，使得

$$\langle q_0^*, y_0\rangle < \alpha\langle q_0^*, p\rangle + \delta. \qquad (7.1.9)$$

考虑 (7.1.8) + (7.1.9)，有

$$\langle q_0^*, y^0\rangle + \langle q_0^*, \bar{y}^0\rangle < 0,$$

但它与(7.1.6)矛盾. 另外，由 $x^0 \neq \mathbf{0}$，故存在 $x_0^* \in \mathscr{X}^*$ 使得 $\langle x_0^*, x^0\rangle = \alpha$，即有

$$\inf\langle q_0^*, \psi(x^0)\rangle = \langle x_0^*, x^0\rangle\langle q_0^*, p\rangle.$$

从(7.1.5)和(7.1.7)可以推证下面两式成立：

$$\langle x_0^*, x^0\rangle p - y \notin \mathrm{int}K \quad \forall y \in \psi(x^0), \qquad (7.1.10)$$

$$-\langle x_0^*, x^0\rangle p - y \notin \mathrm{int}K \quad \forall y \in \psi(-x^0). \qquad (7.1.11)$$

因为假若(7.1.10)不成立，则存在 $y' \in \psi(x^0)$ 使

$$\langle x_0^*, x^0\rangle p - y' \in \mathrm{int}K,$$

从 $q_0^* \in K^* \setminus \{\mathbf{0}\}$ 则得

$$\langle q_0^*, \langle x_0^*, x^0\rangle p\rangle - \langle q_0^*, y'\rangle > 0,$$

这与(7.1.5)矛盾. 类似可证(7.1.11). 现设 $\mathscr{L} = \{x \in \mathscr{X} \mid x = \alpha x^0, \alpha \in R\}$，显然 $\mathscr{L}$ 是 $\mathscr{X}$ 中的非空线性子空间. 我们证明有

$$\mathbf{0} \in \mathscr{E}_w\Big[\bigcup_{x\in\mathscr{L}}\big[\psi(x) - \langle x_0^*, x\rangle p\big], K\Big]. \qquad (7.1.12)$$

事实上，假若 $\mathbf{0} \notin \mathscr{E}_w\Big[\bigcup_{x\in\mathscr{L}}\big[\psi(x) - \langle x_0^*, x\rangle p\big], K\Big]$，则存在 $\alpha' \in R$ 使得

$$\langle x_0^*, \alpha' x^0\rangle p - y \in \mathrm{int}K, \quad y \in \psi(\alpha' x^0). \qquad (7.1.13)$$

下面分三种情形讨论：

(a) 当 $\alpha' = 0$ 时，从(7.1.13)有 $-y \in \text{int}K$，$y \in \psi(\boldsymbol{0})$. 但由已知 $\psi(\boldsymbol{0}) \subset K$ 和 K 是尖锥，故有 $-y \notin K$，这导致矛盾.

(b) 当 $\alpha' > 0$ 时，从(7.1.13)和 $\psi(\alpha' x^0) = \alpha'\psi(x^0)$ 可得

$$\langle x_0^*, x^0\rangle p - \frac{y}{\alpha'} \in \text{int}K, \quad \frac{y}{\alpha'} \in \psi(x^0),$$

它与(7.1.10)相矛盾.

(c) 当 $\alpha' < 0$ 时，从(7.1.13)和 $\psi(\alpha' x^0) = -\alpha'\psi(-x^0)$ 得

$$-\langle x_0^*, x^0\rangle p - \frac{y}{-\alpha'} \in \text{int}K, \quad \frac{y}{-\alpha'} \in \psi(-x^0),$$

它与(7.1.11)相矛盾.

现在考虑 $x^0 = \boldsymbol{0}$的情况. 设 $\mathscr{L} = \{\boldsymbol{0}\}$，由于 $\boldsymbol{0} \in \psi(\boldsymbol{0}) \subset K$，故 $-\psi(\boldsymbol{0}) \notin \text{int}K$. 对 $x_0^* \in \mathscr{X}^*$ 有

$$\boldsymbol{0} \in \mathscr{E}_w\Big(\bigcup_{x\in\mathscr{L}}\big[\psi(x) - \langle x_0^*, x\rangle p\big], K\Big),$$

此时显然有(7.1.4)成立. 根据定理 7.1.2 和(7.1.13)便知，存在 $x^* \in \mathscr{X}^*$ 使得

$$\langle x^*, x^0\rangle = \langle x_0^*, x^0\rangle$$

和 $\boldsymbol{0} \in \mathscr{E}_w\Big(\bigcup_{x\in\mathscr{X}}\big[\psi(x) - \langle x_0^*, x\rangle p\big], K\Big)$. 定理证毕. □

由定理 7.1.5 易得下面两个推论.

推论 7.1.6　所有假设与定理 7.1.5 的相同，则存在 $x^* \in \mathscr{X}^*$ 和 $q^* \in K^* \setminus \{\boldsymbol{0}\}$，使得

$$\inf\langle q_0^*, \psi(x^0)\rangle = \langle x^*, x^0\rangle\langle q_0^*, p\rangle$$

和

$$\langle q^*, y\rangle \geqslant \langle x^*, x\rangle\langle q^*, p\rangle \quad \forall x \in \mathscr{X}, y \in \psi(x).$$

推论 7.1.7　设实值函数 $f: \mathscr{X} \to R$ 是连续的，并且在 $\mathscr{X}$ 上是 R_+-次线性的，$p = 1$，$f(\boldsymbol{0}) = 0$. 若存在 x^0 使得 $f(x^0) > -\infty$，则存在线性泛函 $x^* \in \mathscr{X}^*$，使得 $f(x^0) = \langle x^*, x^0\rangle$ 和

$$\langle x^*, x\rangle \leqslant f(x) \quad \forall x \in \mathscr{X}.$$

推论 7.1.7 就是通常函数(泛函)的 Hahn-Banach 定理. 为了区别,我们将定理 7.1.2 和定理 7.1.5 称为有效 Hahn-Banach 定理.

§7.2 C-锥弱次微分

本节从 C-切锥来定义集值映射的 K-上图象的 C-切锥. 由此,引进 C-锥弱方向导数和 C-锥弱次微分.

设 $\mathscr{X}$ 和 $\mathscr{Y}$ 是局部凸线性拓扑空间, $x \in \mathscr{X}$, $\psi: \mathscr{X} \to 2^{\mathscr{Y}}$, $x \mapsto \psi(x)$ 是集值映射,又 $K \subset \mathscr{Y}$ 是内部非空的尖闭凸锥. 由注 5.2.3 知, ψ 在 $\mathscr{X}$ 上的 K-上图象 $K\text{-epi}\,\psi$ 在点 $(x, y) \in K\text{-epi}\,\psi$ 处的 C-切锥 $T_{K\text{-epi}\psi}(x, y)$ 具有如下性质:

$(d, r) \in T^C_{K\text{-epi}\psi}(x, y)$ 当且仅当对任何收敛于 (x, y) 的序列 $\{(x^k, y^k)\} \subset K\text{-epi}\,\psi$ 和收敛于 0 的正数列 $\{t_k\}$, 存在一序列 $\{(d^k, r^k)\} \in \mathscr{X} \times \mathscr{Y}$ 收敛于 (d, r), 使得有

$$(x^k + t_k d^k,\ y^k + t_k r^k) \in K\text{-epi}\,\psi,\ k = 1, 2, \cdots,$$

即 $y^k + t_k r^k \in \psi(x^k + t_k d^k) + K$. 由定理 5.2.1 知 $T^C_{K\text{-epi}\psi}(x, y)$ 是闭凸的.

定义 7.2.1 设集合 $\psi(x) \neq \varnothing$ $(\forall x \in \mathscr{X})$, 点 $x^0 \in \mathscr{X}$, $y^0 \in \psi(x^0)$, $d \in \mathscr{X}$. 记集合

$$\Psi(x, y, d) = \{r \in \mathscr{Y} \mid (d, r) \in T^C_{K\text{-epi}\psi}(x, y)\}. \tag{7.2.1}$$

若 $y' \in \Psi(x, y, d)$, 并且不存在 $y \in \Psi(x, y, d)$, 使得

$$y' - y \in \operatorname{int} K,$$

则称 y' 是集值映射 ψ 在点 (x^0, y^0) 处沿方向 d 的 *CK-弱方向导数*. ψ 在点 (x^0, y^0) 处沿 d 的所有 CK-弱方向导数组成的集合记作

$\psi_w^C(x^0, y^0; d)$,即 $\psi_w^C(x^0, y^0; d) = \mathscr{E}_w\big(\Psi(x^0, y^0, d), K\big)$.

注 7.2.1　设 $\mathscr{B}$ 是 Banach 空间,$f: \mathscr{B} \to R \cup \{+\infty\}$ 是广义实值函数,并且 $K = R_+$, $x^0 \in \mathscr{B}$, $y^0 = f(x^0)$, $d \in \mathscr{B}$. 由定义 7.2.1 我们有

$$f_w^C(x^0, y^0; d) = f^C(x^0; d) = \inf\{\eta \mid \eta \in \Psi(x^0, y^0, d)\},$$

其中 $f^C(x^0; d)$是 f 的 C-切导数. 根据定理 5.4.2,当 f 在 x^0 处是 Lipschitz 的,则有

$$f_w^C(x^0, y^0; d) = \limsup_{\substack{x \to x^0 \\ t \to 0^+}} \frac{f(x + td) - f(x)}{t} = f^0(x^0; d).$$

可见,CK-弱方向导数是一般函数的 G-方向导数的推广. 记

$$\hat{D} = \{d \in \mathscr{X} \mid \Psi(x^0, y^0, d) \neq \varnothing\},$$

$$K\text{-epi}\,\Psi = \{(x, y) \mid x \in \mathscr{Y}, y \in \Psi(x^0, y^0, x) + K\}.$$

定义 7.2.2　设点 $x^0 \in \mathscr{X}$, $y^0 \in \psi(x^0)$, $p \in \operatorname{int}K$, 则称集合

$$\partial_w^C \psi(x^0, y^0)_p = \Big\{x^* \in \mathscr{X}^* \ \Big|\ 0 \in \mathscr{E}_w\Big(\bigcup_{x \in \hat{D}} \big[\psi_w^C(x^0, y^0; d) - \langle x^*, d\rangle p\big], K\Big)\Big\}$$

是集值映射 ψ 在点(x^0, y^0)处关于 p 的 CK-弱次微分,称 $x^* \in \partial_w^C \psi(x^0, y^0)_p$ 是 ψ 在点(x^0, y^0)处关于 p 的 CK-弱次梯度. 若 $\partial_w^C \psi(x^0, y^0)_p \neq \varnothing$, 则称 ψ 在点(x^0, y^0)处关于 p 是CK-弱次可微的.

定理 7.2.1　设 $x \in \mathscr{X}$, $y \in \psi(x)$, $\hat{D} = \mathscr{X}$. 若 $\Psi(x, y, 0)$ 是 K-有界的,则

$$\psi_w^C(x, y; d) \neq \varnothing \quad \forall d \in \mathscr{X},$$

$$\partial_w^C \psi(x, y)_p = \Big\{x^* \in \mathscr{X}^* \ \Big|\ 0 \in \mathscr{E}_w\Big(\bigcup_{d \in \mathscr{X}} \big[\Psi(x, y, d) - \langle x^*, d\rangle p\big], K\Big)\Big\}$$

$$= \left\{ \boldsymbol{x}^* \in \mathscr{X}^* \ \middle|\ \boldsymbol{0} \in \mathscr{E}_w\left(\bigcup_{(\boldsymbol{d}, \boldsymbol{r}) \in T^C_{K\text{-epi}\varphi}(\boldsymbol{x}, \boldsymbol{y})} \left[\boldsymbol{r} - \langle \boldsymbol{x}^*, \boldsymbol{d}\rangle \boldsymbol{p}\right], K\right) \right\}.$$

证明 取任意的 $\boldsymbol{d}^1, \boldsymbol{d}^2 \in \mathscr{X}$，任意的 $\boldsymbol{y}^1 \in \Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}^1)$ 和任意的 $\boldsymbol{y}^2 \in \Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}^2)$，由定义 7.2.1 可知 $(\boldsymbol{d}^1, \boldsymbol{y}^1), (\boldsymbol{d}^2, \boldsymbol{y}^2) \in T^C_{K\text{-epi}\varphi}(\boldsymbol{x}, \boldsymbol{y})$. 因为 $T^C_{K\text{-epi}\varphi}(\boldsymbol{x}, \boldsymbol{y})$ 是凸锥，$(\boldsymbol{d}^1 + \boldsymbol{d}^2, \boldsymbol{y}^1 + \boldsymbol{y}^2) \in T^C_{K\text{-epi}\varphi}(\boldsymbol{x}, \boldsymbol{y})$，由(7.2.1)我们有 $\boldsymbol{y}^1 + \boldsymbol{y}^2 \in \Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}^1 + \boldsymbol{d}^2)$，因此得

$$\Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}^1) + \Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}^2) \subset \Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}^1 + \boldsymbol{d}^2). \tag{7.2.2}$$

取任意的 $\boldsymbol{d} \in \mathscr{X}$，令 $\boldsymbol{b} \in \Psi(\boldsymbol{x}, \boldsymbol{y}, -\boldsymbol{d})$，根据(7.2.2)有

$$\Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}) + \boldsymbol{b} \subset \Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0}).$$

因为 $\Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0})$ 是 K-有界的，所以存在 $\boldsymbol{a} \in \mathscr{Y}$ 使得 $\Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0}) \subset \boldsymbol{a} + K$，因此

$$\Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}) \subset \boldsymbol{a} - \boldsymbol{b} + K.$$

根据定义 6.5.2 的(i)知 $\Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d})$ 也是 K-有界的. 因为 $T^C_{K\text{-epi}\varphi}(\boldsymbol{x}, \boldsymbol{y})$ 是闭的，K 是闭的，所以 $\Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d})$ 是 K-闭集. 据此，由引理 6.5.1 知 $\Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d})$ 是 K-半紧的. 于是，根据引理 6.2.5 得

$$\varphi^C_w(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}) \neq \varnothing \quad \forall \boldsymbol{d} \in \mathscr{X}.$$

再由引理 6.5.2 便知

$$\Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}) \subset \varphi^C_w(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}) + K. \tag{7.2.3}$$

设集合

$$S_1 = \left\{ \boldsymbol{x}^* \in \mathscr{X}^* \ \middle|\ \boldsymbol{0} \in \mathscr{E}_w\left(\bigcup_{x \in \mathscr{X}} \left[\Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}) - \langle \boldsymbol{x}^*, \boldsymbol{d}\rangle \boldsymbol{p}\right], K\right) \right\},$$

$$S_2 = \left\{ \boldsymbol{x}^* \in \mathscr{X}^* \ \middle|\ \boldsymbol{0} \in \mathscr{E}_w\left(\bigcup_{(\boldsymbol{d}, \boldsymbol{r}) \in T^C_{K\text{-epi}\varphi}(\boldsymbol{x}, \boldsymbol{y})} \left[\boldsymbol{r} - \langle \boldsymbol{x}^*, \boldsymbol{d}\rangle \boldsymbol{p}\right], K\right) \right\}.$$

显然有 $S_1 = S_2$，因此仅需证明 $S_1 = \partial_w^C \psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$. 若 $\boldsymbol{x}^* \in S_1$，则不存在 $\boldsymbol{y}' \in \Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d})$ 使得

$$\boldsymbol{0} - \left(\boldsymbol{y}' - \langle \boldsymbol{x}^*, \boldsymbol{d} \rangle \boldsymbol{p}\right) \in \operatorname{int} K.$$

由此，按定义 7.2.1 知 $\psi_w^C(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}) \subset \Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d})$，于是不存在 $\boldsymbol{y}'' \in \psi_w^C(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d})$ 使

$$\boldsymbol{0} - \left(\boldsymbol{y}'' - \langle \boldsymbol{x}^*, \boldsymbol{d} \rangle \boldsymbol{p}\right) \in \operatorname{int} K.$$

因此，由定义 7.2.2 得知 $\boldsymbol{x}^* \in \partial_w^C \psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$.

现设 $\boldsymbol{x}^* \in \partial_w^C \psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$，反之假若 $\boldsymbol{x}^* \notin S_1$，则存在 $\overline{\boldsymbol{y}} \in \Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d})$，使得

$$\boldsymbol{0} - \left(\overline{\boldsymbol{y}} - \langle \boldsymbol{x}^*, \boldsymbol{d} \rangle \boldsymbol{p}\right) \in \operatorname{int} K.$$

据此，由(7.2.3)，存在 $\boldsymbol{q}' \in K$ 和 $\overline{\boldsymbol{y}}' \in \psi_w^C(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d})$，使得 $\overline{\boldsymbol{y}} = \overline{\boldsymbol{y}}' + \boldsymbol{q}'$，从而

$$\boldsymbol{0} - \left(\overline{\boldsymbol{y}}' + \boldsymbol{q}' - \langle \boldsymbol{x}^*, \boldsymbol{d} \rangle \boldsymbol{p}\right) \in \operatorname{int} K,$$

即 $\boldsymbol{0} - \left(\overline{\boldsymbol{y}}' - \langle \boldsymbol{x}^*, \boldsymbol{d} \rangle \boldsymbol{p}\right) \in \operatorname{int} K$. 于是由定义 7.2.2 可知 $\boldsymbol{x}^* \notin \partial_w^C \psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$，这导致与已设 $\boldsymbol{x}^* \in \partial_w^C \psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$ 矛盾. □

下面给出锥弱次微分的存在性定理.

定理 7.2.2 设 $\boldsymbol{x} \in \mathscr{X}$，$\boldsymbol{y} \in \psi(\boldsymbol{x})$，$\hat{D} = \mathscr{X}$，$\Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0}) \subset K$，$\operatorname{int}(K\text{-epi}\Psi) \neq \varnothing$（或 $\Psi(\boldsymbol{x}, \boldsymbol{y}, \cdot)$ 在某个 $\overline{\boldsymbol{d}} \in \mathscr{X}$ 处是 $(-K)$-有界的）. 若存在 $\boldsymbol{d}^0 \in \mathscr{X}$ 和 $\boldsymbol{q}_0^* \in K^* \backslash \{\boldsymbol{0}\}$ 使得 $\inf \langle \boldsymbol{q}_0^*, \Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}^0) \rangle > -\infty$，则存在线性泛函 $\boldsymbol{x}^* \in \mathscr{X}^*$ 使得

$$\inf \langle \boldsymbol{q}_0^*, \Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}^0) \rangle = \langle \boldsymbol{x}^*, \boldsymbol{d}^0 \rangle \langle \boldsymbol{q}_0^*, \boldsymbol{p} \rangle,$$

并且 $\boldsymbol{x}^* \in \partial_w^C \psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$.

证明 因为 $T_{K\text{-epi}\psi}^C(\boldsymbol{x}, \boldsymbol{y})$ 是凸锥，所以对任意的 $\lambda > 0$ 和 $(\boldsymbol{d}, \boldsymbol{r}) \in T_{K\text{-epi}\psi}^C(\boldsymbol{x}, \boldsymbol{y})$ 有 $(\lambda \boldsymbol{d}, \lambda \boldsymbol{r}) \in T_{K\text{-epi}\psi}^C(\boldsymbol{x}, \boldsymbol{y})$. 由(7.2.1)知 $\lambda \boldsymbol{r} \in \Psi(\boldsymbol{x}, \boldsymbol{y}, \lambda \boldsymbol{d})$，故 $\Psi(\boldsymbol{x}, \boldsymbol{y}, \lambda \boldsymbol{d}) = \lambda \Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d})$. 显然，我们有

$$\Psi(x, y, d^1) + \Psi(x, y, d^2) \subset \Psi(x, y, d^1 + d^2)$$

$$\forall d^1, d^2 \in \mathscr{X}.$$

因此,$\Psi(x, y, d)$关于d是K-次线性的.又显然有$0 \in \Psi(x, y, 0)$,于是由定理 7.1.5 和定理 7.2.1 得到结论.▯

定理 7.2.3 设$x \in \mathscr{X}$, $y \in \psi(x)$, $\hat{D} = \mathscr{X}$. 若$\Psi(x, y, 0)$是K-有界的,则$x^* \in \partial_w^C \psi(x, y)_p$当且仅当存在$q^* \in K^* \setminus \{0\}$,使得

$$\langle q^*, y' \rangle \geqslant \langle x^*, d \rangle \langle q^*, p \rangle \quad \forall d \in \mathscr{X}, \forall y' \in \psi_w^C(x, y; d). \tag{7.2.4}$$

证明 必要性. 设$x^* \in \partial_w^C \psi(x, y)_p$,作集合

$$E = \{y \in \mathscr{Y}^* \mid y = -r + \langle x^*, d \rangle p - q,$$

$$(d, r) \in T_{K\text{-epi}\psi}^C(x, y), q \in K\}.$$

易知E是凸集,现在证明$\mathrm{int}K \cap E = \varnothing$. 反之,假若有$q' \in \mathrm{int}K \cap E$,即$-r + \langle x^*, d \rangle p - q = q'$,则有$-r + \langle x^*, d \rangle p = q + q' \in \mathrm{int}K$. 由定理 7.2.1 得$x^* \notin \partial_w^C \psi(x, y)_p$,导致与已设$x^* \in \partial_w^C \psi(x, y)_p$矛盾. 据此,由定理 1.3.7 可知,存在$q^* \in \mathscr{Y}^* \setminus \{0\}$使得

$$\langle q^*, q \rangle \geqslant \langle q^*, -r + \langle x^*, d \rangle p - q' \rangle$$

$$\forall q \in K, \forall (d, r) \in T_{K\text{-epi}\psi}^C(x, y), q' \in K. \tag{7.2.5}$$

在(7.2.5)中令$r = 0$, $d = 0$, $q' = 0$得$\langle q^*, q \rangle \geqslant 0$ ($\forall q \in K$),故有$q^* \in K^* \setminus \{0\}$. 在(7.2.5)中再令$q' = 0$, $q = 0$,得到

$$\langle q^*, r \rangle \geqslant \langle x^*, d \rangle \langle q^*, p \rangle \quad \forall (d, r) \in T_{K\text{-epi}\psi}^C(x, y).$$

由定义 7.1.1 和$\psi_w^C(x, y; d) \subset \Psi(x, y, d)$即得(7.2.4).

充分性. 现设(7.2.4)成立,反之,假设$x^* \notin \partial_w^C \psi(x, y)_p$. 由定义 7.2.2 知存在$y' \in \psi_w^C(x, y; d)$,使得

$$\boldsymbol{0}-\left(\boldsymbol{y}'-\langle\boldsymbol{x}^*,\boldsymbol{d}\rangle\boldsymbol{p}\right)\in \mathrm{int}K.$$

由 $\boldsymbol{q}^*\in K^*\backslash\{\boldsymbol{0}\}$ 有

$$\langle\boldsymbol{q}^*,-\boldsymbol{y}'+\langle\boldsymbol{x}^*,\boldsymbol{d}\rangle\boldsymbol{p}\rangle>0,$$

这导致与(7.2.4)矛盾.▯

推论 7.2.4　设 $\boldsymbol{x}\in\mathscr{X}$, $\boldsymbol{y}\in\psi(\boldsymbol{x})$, $\hat{D}=\mathscr{X}$, $\boldsymbol{p}^1$, $\boldsymbol{p}^2\in\mathrm{int}K$. 若$\Psi(\boldsymbol{x},\boldsymbol{y},\boldsymbol{0})$是 K-有界的,并且 $\boldsymbol{x}^*\in\partial_w^C\psi(\boldsymbol{x},\boldsymbol{y})_{\boldsymbol{p}^1}$, 则存在 $\boldsymbol{q}^*\in K^*\backslash\{\boldsymbol{0}\}$, 使得 $\alpha\boldsymbol{x}^*\in\partial_w^C\psi(\boldsymbol{x},\boldsymbol{y})_{\boldsymbol{p}^2}$, 其中 $\alpha=\langle\boldsymbol{q}^*,\boldsymbol{p}^1\rangle/\langle\boldsymbol{q}^*,\boldsymbol{p}^2\rangle$.

证明　由定理 7.2.3 容易推得证明.▯

推论 7.2.4 说明了 ψ 的 K-次梯度与 $\boldsymbol{p}$ 无关.

定理 7.2.5　设 $\boldsymbol{x}\in\mathscr{X}$, $\boldsymbol{y}\in\psi(\boldsymbol{x})$, $\hat{D}=\mathscr{X}$. 若$\Psi(\boldsymbol{x},\boldsymbol{y},\boldsymbol{0})$是 K-有界的,并且 $\partial_w^C\psi(\boldsymbol{x},\boldsymbol{y})_{\boldsymbol{p}}\neq\varnothing$, 则 $\partial_w^C\psi(\boldsymbol{x},\boldsymbol{y})_{\boldsymbol{p}}$ 是闭凸集.

证明　先证闭性. 用反证法,假设有序列 $\{\boldsymbol{x}_k^*\}\subset\partial_w^C\psi(\boldsymbol{x},\boldsymbol{y})_{\boldsymbol{p}}$, $\boldsymbol{x}_k^*\to\boldsymbol{x}^*$, 但 $\boldsymbol{x}^*\notin\partial_w^C\psi(\boldsymbol{x},\boldsymbol{y})_{\boldsymbol{p}}$, 则由定理 7.2.1 知,存在 $(\boldsymbol{d},\boldsymbol{r})\in T_{K\text{-epi}\psi}^C(\boldsymbol{x},\boldsymbol{y})$ 使得

$$-\left(\boldsymbol{r}-\langle\boldsymbol{x}^*,\boldsymbol{d}\rangle\boldsymbol{p}\right)\in\mathrm{int}K.$$

因为 $-\boldsymbol{r}+\langle\boldsymbol{x}_k^*,\boldsymbol{d}\rangle\boldsymbol{p}\to-\boldsymbol{r}+\langle\boldsymbol{x}^*,\boldsymbol{d}\rangle\boldsymbol{p}(k\to\infty)$, 所以存在充分大的 k',使得

$$-\left(\boldsymbol{r}-\langle\boldsymbol{x}_{k'}^*,\boldsymbol{d}\rangle\boldsymbol{p}\right)\in\mathrm{int}K.$$

再由定理 7.2.1 得知 $\boldsymbol{x}_{k'}^*\notin\partial_w^C\psi(\boldsymbol{x},\boldsymbol{y})_{\boldsymbol{p}}$, 这导致与假设矛盾.

现证凸性. 设 $\boldsymbol{x}_1^*$, $\boldsymbol{x}_2^*\in\partial_w^C\psi(\boldsymbol{x},\boldsymbol{y})_{\boldsymbol{p}}$, 由定理 7.2.3 可知,存在 $\boldsymbol{q}_1^*$, $\boldsymbol{q}_2^*\in K^*\backslash\{\boldsymbol{0}\}$ 使得

$$\langle\boldsymbol{q}_1^*,\boldsymbol{r}\rangle\geqslant\langle\boldsymbol{x}_1^*,\boldsymbol{d}\rangle\langle\boldsymbol{q}_1^*,\boldsymbol{p}\rangle\quad\forall(\boldsymbol{d},\boldsymbol{r})\in T_{K\text{-epi}\psi}^C(\boldsymbol{x},\boldsymbol{y}),\tag{7.2.6}$$

$$\langle\boldsymbol{q}_2^*,\boldsymbol{r}\rangle\geqslant\langle\boldsymbol{x}_2^*,\boldsymbol{d}\rangle\langle\boldsymbol{q}_2^*,\boldsymbol{p}\rangle\quad\forall(\boldsymbol{d},\boldsymbol{r})\in T_{K\text{-epi}\psi}^C(\boldsymbol{x},\boldsymbol{y}).\tag{7.2.7}$$

对任意的 $\lambda \in (0, 1)$，设 $\boldsymbol{x}_\lambda^* = \lambda \boldsymbol{x}_1^* + (1-\lambda)\boldsymbol{x}_2^*$，我们要证明 $\boldsymbol{x}_\lambda^* \in \partial_w^C \psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$. 反之，假设 $\boldsymbol{x}_\lambda^* \notin \partial_w^C \psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$，则存在 $(\boldsymbol{d}, \boldsymbol{r}) \in T_{K\text{-epi}\psi}^C(\boldsymbol{x}, \boldsymbol{y})$ 使得 $-\boldsymbol{r} + \langle \boldsymbol{x}_\lambda^*, \boldsymbol{d}\rangle \boldsymbol{p} \in \mathrm{int}K$. 由 $\boldsymbol{q}_1^*, \boldsymbol{q}_2^* \in K^* \backslash \{\boldsymbol{0}\}$ 和定理 1.4.11，可推得

$$\left\langle \boldsymbol{q}_1^*, -\boldsymbol{r} + \langle \boldsymbol{x}_\lambda^*, \boldsymbol{d}\rangle \boldsymbol{p} \right\rangle > 0, \tag{7.2.8}$$

$$\left\langle \boldsymbol{q}_2^*, -\boldsymbol{r} + \langle \boldsymbol{x}_\lambda^*, \boldsymbol{d}\rangle \boldsymbol{p} \right\rangle > 0. \tag{7.2.9}$$

考虑 (7.2.6) + (7.2.8) 和 (7.2.7) + (7.2.9)，可得

$$\langle \boldsymbol{x}_\lambda^*, \boldsymbol{d}\rangle > \langle \boldsymbol{x}_1^*, \boldsymbol{d}\rangle, \tag{7.2.10}$$

$$\langle \boldsymbol{x}_\lambda^*, \boldsymbol{d}\rangle > \langle \boldsymbol{x}_2^*, \boldsymbol{d}\rangle. \tag{7.2.11}$$

从 λ(7.2.10) + $(1-\lambda)$ × (7.2.11)，则得

$$\langle \boldsymbol{x}_\lambda^*, \boldsymbol{d}\rangle > \langle \boldsymbol{x}_\lambda^*, \boldsymbol{d}\rangle,$$

导致矛盾. ▯

定理 7.2.6 设 $\boldsymbol{x} \in \mathscr{X}$，$\boldsymbol{y} \in \psi(\boldsymbol{x})$，$\hat{D} = \mathscr{X}$，$\Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0}) \subset K$. 若 $\mathrm{int}(K - \mathrm{epi}\Psi) \neq \varnothing$，$\partial_w^C(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}} \neq \varnothing$，则对任给的 $\boldsymbol{d} \in \mathscr{X}$ 存在 $\boldsymbol{q}^* \in K^* \backslash \{\boldsymbol{0}\}$，使得

$$\begin{aligned} &\inf \left\langle \boldsymbol{q}^*, \psi_w^C(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}) \right\rangle \\ &= \max \left\{ \langle \boldsymbol{x}^*, \boldsymbol{d}\rangle \langle \boldsymbol{q}^*, \boldsymbol{p}\rangle \,\middle|\, \boldsymbol{x}^* \in \partial_w^C \psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}} \right\}. \end{aligned}$$

证明 任给 $\boldsymbol{d}^1 \in \mathscr{X}$，设 $\overline{\boldsymbol{x}}^* \in \partial_w^C \psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$，使得

$$\langle \overline{\boldsymbol{x}}^*, \boldsymbol{d}^1\rangle \geqslant \langle \boldsymbol{x}^*, \boldsymbol{d}^1\rangle \quad \forall \boldsymbol{x}^* \in \partial_w^C \psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}. \tag{7.2.12}$$

由定理 7.2.3，对于 $\overline{\boldsymbol{x}}^* \in \partial_w^C \psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$ 存在 $\boldsymbol{q}^* \in K^* \backslash \{\boldsymbol{0}\}$ 使

$$\langle \boldsymbol{q}^*, \boldsymbol{y}'\rangle \geqslant \langle \overline{\boldsymbol{x}}^*, \boldsymbol{d}\rangle \langle \boldsymbol{q}^*, \boldsymbol{p}\rangle \quad \forall\, \boldsymbol{d} \in \mathscr{X}, \forall \boldsymbol{y}' \in \psi_w^C(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}).$$

从上式显然有

$$\inf \left\langle \boldsymbol{q}^*, \psi_w^C(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}^1) \right\rangle \geqslant \langle \overline{\boldsymbol{x}}^*, \boldsymbol{d}^1\rangle \langle \boldsymbol{q}^*, \boldsymbol{p}\rangle.$$

于是，由定理 7.2.2，存在 $x_1^* \in \partial_w^C\psi(x, y)_p$ 使得

$$\inf\langle q^*, \psi_w^C(x, y; d^1)\rangle = \langle x_1^*, d^1\rangle\langle q^*, p\rangle,$$

因此有 $\langle x_1^*, d^1\rangle \geqslant \langle \bar{x}^*, d^1\rangle$. 再由(7.2.12)得

$$\langle x_1^*, d^1\rangle = \langle \bar{x}^*, d^1\rangle,$$

即定理结论成立. □

推论 7.2.7　设 $\mathscr{B}$ 是 Banach 空间. 若实值函数 $f: \mathscr{B} \to R$ 在 $x \in \mathscr{B}$ 处是 Lipschitz 的，$y = f(x)$，$p = 1$，$K = R_+$，则

$$f^0(x, y; d) = \max\{\langle x^*, d\rangle \mid x^* \in \partial^0 f(x)\} \quad \forall d \in \mathscr{B}.$$

证明　由注 7.2.1 和定理 7.2.6，即知结论成立. □

现在，讨论锥凸集值映射的锥弱次微分与第 6 章的锥次微分之间的关系.

定理 7.2.8　设 $x \in \mathscr{X}$，$y \in \psi(x)$，集值映射 ψ 在 $\mathscr{X}$ 上是 K-凸的，$\hat{D} = \mathscr{X}$. 若 $\Psi(x, y, \theta)$ 是 K-有界的，则

$$\partial_w^C\psi(x, y)_p \supset \Big\{x^* \in \mathscr{X}^* \,\Big|\, y - \langle x^*, x\rangle p \in \mathscr{E}_w\Big(\bigcup_{x \in \mathscr{X}}[\psi(x') - \langle x^*, x'\rangle p], K\Big)\Big\}.$$

当 $\mathscr{X}$ 和 $\mathscr{Y}$ 是 Banach 空间时，上式成为等式.

证明　由 ψ 是 K-凸的，集合

$$C_{K\text{-epi}\psi}(x, y) = \mathrm{cl}\Big(\bigcup_{\lambda > 0}\lambda[K\text{-epi}\psi - (x, y)]\Big)$$

非空，由定义 5.2.2 知

$$T_{K\text{-epi}\psi}^C(x, y) \subset C_{K\text{-epi}\psi}(x, y).$$

若 $\mathscr{X}$ 和 $\mathscr{Y}$ 是 Banach 空间，则由定理 5.2.2 知上式为等式. 由此，根据定理 6.5.3 和定义 7.2.2 可得知本定理的结论成立. □

定义 7.2.3　设点 $x^0 \in \mathscr{X}$，$y^0 \in \psi(x^0)$. 若存在 x^0 的邻域 $U(x^0)$ 使得

$$y^0 \in \mathscr{E}_w\big(\psi\big(U(x^0)\big), K\big),$$

即不存在 $x \in U(x^0)$ 和 $y \in \psi(x)$ 使得

$$y^0 - y \in \mathrm{int}K,$$

则称点(x^0, y^0)是集值映射 ψ 的局部弱有效解.

定理 7.2.9　设 $x \in \mathscr{X}$, $y \in \psi(x)$, $\hat{D} = \mathscr{X}$. 若$\Psi(x, y, \boldsymbol{0})$是 K-有界的,(x, y)是 ψ 的局部弱有效解,则 $\boldsymbol{0} \in \partial_w^C \psi(x, y)_p$.

证明　用反证法:假设 $\boldsymbol{0} \notin \partial_w^C \psi(x, y)_p$. 由定理 7.2.1 可知,存在 $r \in \Psi(x, y, d)$ 使得

$$-r + \langle \boldsymbol{0}, d \rangle p \in \mathrm{int}K. \tag{7.2.13}$$

由此,对任何收敛于$(x, y, 0)$的序列$\{(x^k, y, t_k)\}$存在收敛于(d, r)的序列$\{(d^k, r^k)\}$,使得

$$y + t_k r^k \in \psi(x^k + t_k d^k) + K, \ k = 1, 2, \cdots.$$

于是,存在 $y^k \in \psi(x^k + t_k d^k)$ 和 $q^k \in K$,使得

$$y + t_k r^k = y^k + q^k, \ k = 1, 2, \cdots,$$

从而有 $r^k = \dfrac{y^k - q^k - y}{t_k} \to r$. 又因为 $x^k + t_k r^k \to x$,故存在充分大的 k',使 $x^{k'} + t_{k'} r^{k'} \in U(x)$, $y^{k'} \in \psi(x^{k'} + t_{k'} r^{k'})$,以及$-\dfrac{y^{k'} - q^{k'} - y}{t^{k'}} \in \mathrm{int}K$,或即

$$y - y^{k'} \in \mathrm{int}K.$$

据此,由定义 7.2.3 得到 $y \notin \mathscr{E}_w(U(x), K)$,这与已知矛盾. 定理得证. □

§ 7.3　G-锥弱次微分

本节先介绍一种 G-切锥,它比 C-切锥小. 从这种 G-切锥出

发，我们引进集值映射的 G-锥弱方向导数和 G-锥弱次微分.

设 $\mathscr{X}$ 和 $\mathscr{Y}$ 是局部凸线性拓扑空间，$\psi: \mathscr{X} \to 2^{\mathscr{Y}}$，$\boldsymbol{x} \mapsto \psi(\boldsymbol{x})$ 是集值映射，$K \subset \mathscr{Y}$ 是内部非空的尖闭凸锥.

定义 7.3.1 设点 $\boldsymbol{x}^0 \in \mathscr{X}$，$\boldsymbol{y}^0 \in \mathscr{Y}$. 我们称集合 $G_{K\text{-epi}\psi}(\boldsymbol{x}^0, \boldsymbol{y}^0)$ 是实值映射 ψ 在 $\mathscr{X}$ 上的 K-上图象 K-epiψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的 G-切锥，即对任意的 $(\boldsymbol{d}, \boldsymbol{r}) \in G_{K\text{-epi}\psi}(\boldsymbol{x}^0, \boldsymbol{y}^0) \subset \mathscr{X} \times \mathscr{Y}$，若对任何收敛于$(\boldsymbol{x}^0, \boldsymbol{y}^0)$的序列 $\{(\boldsymbol{x}^k, \boldsymbol{y}^k)\} \subset K$-epi$\psi$ 和收敛于 0 的正数列 $\{t_k\}(k = 1, 2, \cdots)$，存在一个收敛于 $\boldsymbol{r}$ 的序列 $\{\boldsymbol{r}^k\} \subset \mathscr{Y}$，使得

$$(\boldsymbol{x}^k + t_k\boldsymbol{d}, \boldsymbol{y}^k + t_k\boldsymbol{r}^k) \in K\text{-epi}\psi, \ k = 1, 2, \cdots.$$

关于 G-切锥有如下性质.

定理 7.3.1 设 $\boldsymbol{x} \in \mathscr{X}$，$\boldsymbol{y} \in \psi(\boldsymbol{x})$.

(1) $(\boldsymbol{0}, \boldsymbol{0}) \in G_{K\text{-epi}\psi}(\boldsymbol{x}, \boldsymbol{y}) \subset T^{C}_{K\text{-epi}\psi}(\boldsymbol{x}, \boldsymbol{y})$.

(2) $G_{K\text{-epi}\psi}(\boldsymbol{x}, \boldsymbol{y})$ 是凸锥.

(3) 若对任何 $(\boldsymbol{d}, \boldsymbol{r}^k) \in G_{K\text{-epi}\psi}(\boldsymbol{x}, \boldsymbol{y})(k = 1, 2, \cdots)$ 有 $(\boldsymbol{d}, \boldsymbol{r}^k) \to (\boldsymbol{d}, \boldsymbol{r})$，则 $(\boldsymbol{d}, \boldsymbol{r}) \in G_{K\text{-epi}\psi}(\boldsymbol{x}, \boldsymbol{y})$.

证明 (1) 由定义 7.3.1 容易直接推得.

(2) 设 $(\boldsymbol{d}, \boldsymbol{r}) \in G_{K\text{-epi}\psi}(\boldsymbol{x}, \boldsymbol{y})$，$\lambda > 0$，按定义 7.3.1，则对任何收敛于$(\boldsymbol{x}, \boldsymbol{y})$的序列 $\{(\boldsymbol{x}^k, \boldsymbol{y}^k)\} \subset K$-epi$\psi$ 和收敛于 0 的正数列 $\{t_k\}$，有收敛于 $\boldsymbol{r}$ 的序列 $\{\boldsymbol{r}^k\} \subset \mathscr{Y}$，使得

$$(\boldsymbol{x}^k + \lambda t_k\boldsymbol{d}, \boldsymbol{y}^k + \lambda t_k\boldsymbol{r}^k) \in K\text{-epi}\psi, \ k = 1, 2, \cdots.$$

由于 $\lambda\boldsymbol{r}^k \to \lambda\boldsymbol{r}$，据定义 7.3.1 得

$$(\lambda\boldsymbol{d}, \lambda\boldsymbol{r}) \in G_{K\text{-epi}\psi}(\boldsymbol{x}, \boldsymbol{y}),$$

因此 $G_{K\text{-epi}\psi}(\boldsymbol{x}, \boldsymbol{y})$ 是一个锥.

为证 $G_{K\text{-epi}\psi}(\boldsymbol{x}, \boldsymbol{y})$ 是凸的，设 $(\boldsymbol{d}^1, \boldsymbol{r}^1), (\boldsymbol{d}^2, \boldsymbol{r}^2) \in G_{K\text{-epi}\psi}(\boldsymbol{x}, \boldsymbol{y})$. 对任何收敛于$(\boldsymbol{x}^k, \boldsymbol{y}^k, 0)$的序列 $\{(\boldsymbol{x}^k, \boldsymbol{y}^k, t_k)\} \subset K$-epi$\psi \times (0, +\infty)$，则存在收敛于 $\boldsymbol{r}^1$ 的序列$\{\boldsymbol{r}^1_k\}$，使得

$$(\boldsymbol{x}^k + t_k\boldsymbol{d}^1, \boldsymbol{y}^k + t_k\boldsymbol{r}^1_k) \in K\text{-epi}\psi, \ k = 1, 2, \cdots.$$

因为 $x^k+t_kd^1\to x$, $y^k+t_kr_k^1\to y$, $t_k\to 0(k\to\infty)$, $(d^2, r^2)\in G_{K\text{-epi}\psi}(x, y)$, 所以存在收敛于 r^2 的序列 $\{r_k^2\}$,使得

$$(x^k+t_kd^1+t_kd^2,\ y^k+t_kr_k^1+t_kr_k^2)\in K\text{-epi}\psi,\ k=1,2,\cdots.$$

由于 $r_k^1+r_k^2\to r^1+r^2$, 故由上式可推得

$$(d^1+d^2,\ r^1+r^2)\in G_{K\text{-epi}\psi}(x, y),$$

于是证明了 $G_{K\text{-epi}\psi}(x, y)$ 是凸的.

(3) 考虑任何收敛于 $(x, y, 0)$ 的序列 $\{(x^k, y^k, t_n)\}\subset K\text{-epi}\psi\times(0,+\infty)$ $(n=1,2,\cdots)$. 由 $(d, r^k)\in G_{K\text{-epi}\psi}(x, y)$ $(k=1,2,\cdots)$, 则对于每一个 $k(=1,2,\cdots)$ 存在序列 $\{r_n^k\}$ 收敛于 $r^k(n\to\infty)$,使得

$$(x^k+t_kd,\ y^k+t_kr_n^k)\in K\text{-epi}\psi,\ n=1,2,\cdots;\ k=1,2,\cdots.$$

设 V 是 $\mathscr{Y}$ 的原点的任意邻域,则存在另一个原点邻域 $U\subset\mathscr{Y}$, 使得 $U+U\subset V$. 因为 $r^k\to r$,所以存在 $N>0$, 使得 $r^k\in r+U$ 对任意的 $k\geqslant N$ 成立. 又因为 $r_n^k\to r^k(n\to\infty)$对于每一个 $k(=1,2,\cdots)$ 成立,故对于每一个 $k(=1,2,\cdots,)$ 均存在 $N_k>0$, 使得

$$r_n^k\in r^k+U,\quad n\geqslant N_k,$$

从而得

$$r_n^k\in r+U+U\subset r+V,\quad n\geqslant N_k,\ k\geqslant N.$$

对每一个 $k(\geqslant N)$ 对应地取一个 $r_n=r_n^k(n=1,2,\cdots)$, 故获得一个序列 $\{r_n\}$ 收敛于 r,并且有

$$(x^n+t_nd,\ y^n+t_nr_n)\in K\text{-epi}\psi,\ n=1,2,\cdots,$$

即 $(d, r)\in G_{K\text{-epi}\psi}(x, y)$. □

定义 7.3.2 设点 $x^0\in\mathscr{X}$, $y^0\in\psi(x^0)$, $d\in\mathscr{X}$. 令集合

$$\Psi^0(x^0, y^0, d)=\{r\in\mathscr{Y}\mid(d, r)\in G_{K\text{-epi}\psi}(x^0, y^0)\},$$

$$\overline{D}=\{d\in\mathscr{X}\mid\Psi^0(x^0, y^0, d)\neq\varnothing\}.$$

若 $y'(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d}) \in \Psi^0(\boldsymbol{x}^0, \boldsymbol{y}^0, \boldsymbol{d})$，并且不存在 $\boldsymbol{y} \in \Psi^0(\boldsymbol{x}^0, \boldsymbol{y}^0, \boldsymbol{d})$ 使得

$$y'(\boldsymbol{x}^0, \boldsymbol{y}^0, \boldsymbol{d}) - \boldsymbol{y} \in \mathrm{int}K,$$

则称 $y'(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d})$是集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处沿方向 $\boldsymbol{d}$ 的 GK-弱方向导数. ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处沿 $\boldsymbol{d}$ 的所有 GK-弱方向导数组成的集合记作 $\psi_w^0(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d})$，即

$$\psi_w^0(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d}) = \mathscr{E}_w\big(\Psi^0(\boldsymbol{x}^0, \boldsymbol{y}^0, \boldsymbol{d}), K\big).$$

定义 7.3.3　设点 $\boldsymbol{x}^0 \in \mathscr{X}$, $\boldsymbol{y}^0 \in \psi(\boldsymbol{x}^0)$, $\overline{D} \neq \varnothing$, $\boldsymbol{p} \in \mathrm{int}K$. 称集合

$$\begin{aligned}&\partial_w^0\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)_{\boldsymbol{p}}\\ &= \left\{\boldsymbol{x}^* \in \mathscr{X}^* \;\middle|\; \boldsymbol{0} \in \mathscr{E}_w\Big(\bigcup_{x \in \overline{D}}\big[\psi_w^0(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d}) - \langle \boldsymbol{x}^*, \boldsymbol{d}\rangle \boldsymbol{p}\big], K\Big)\right\}\end{aligned}$$

是集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处关于 $\boldsymbol{p}$ 的 GK-弱次微分，称 $\boldsymbol{x}^* \in \partial_w^0\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)_{\boldsymbol{p}}$ 是 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处关于 $\boldsymbol{p}$ 的 GK-弱次梯度. 若 $\partial_w^0\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)_{\boldsymbol{p}} \neq \varnothing$，则称 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处关于 $\boldsymbol{p}$ 是 GK-弱次可微的.

定理 7.3.2　设 $\boldsymbol{x} \in \mathscr{X}$, $\boldsymbol{y} \in \psi(\boldsymbol{x})$, $\boldsymbol{d} \in \mathscr{X}$, $\boldsymbol{p} \in \mathrm{int}K$, $\overline{D} = \mathscr{X}$. 若$\Psi^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0})$是 K-有界的，则

$$\psi_w^0(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}) \neq \varnothing \quad \forall \boldsymbol{d} \in \mathscr{X},$$

$$\begin{aligned}&\partial_w^0\psi(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}\\ &= \left\{\boldsymbol{x}^* \in \mathscr{X}^* \;\middle|\; \boldsymbol{0} \in \mathscr{E}_w\Big(\bigcup_{d \in \mathscr{X}}\big[\psi_w^0(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}) - \langle \boldsymbol{x}^*, \boldsymbol{d}\rangle \boldsymbol{p}\big], K\Big)\right\}\\ &= \left\{\boldsymbol{x}^* \in \mathscr{X}^* \;\middle|\; \boldsymbol{0} \in \mathscr{E}_w\Big(\bigcup_{(d, r) \in G_{K\text{-epi}\psi}(x, y)}\big[\boldsymbol{r} - \langle \boldsymbol{x}^*, \boldsymbol{d}\rangle \boldsymbol{p}\big], K\Big)\right\}.\end{aligned}$$

证明　与定理 7.2.1 的证明类似. □

定理 7.3.3　设 $\boldsymbol{x} \in \mathscr{X}$, $\boldsymbol{y} \in \psi(\boldsymbol{x})$, $\boldsymbol{p} \in \mathrm{int}K$, $\overline{D} = \mathscr{X}$. 又设 $\Psi^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0}) \subset K$, $\mathrm{int}(K\text{-epi}\Psi^0) \neq \varnothing$ (或$\Psi^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d})$在某个 $\boldsymbol{d}$ 附近是 K-有界的). 若存在 $\boldsymbol{d}^0 \in \mathscr{X}$ 和 $\boldsymbol{q}_0^* \in K^* \setminus \{\boldsymbol{0}\}$，使得

$\inf\left\langle q_0^*, \Psi^0(x, y; d^0)\right\rangle > -\infty$，则存在线性泛函 $x^* \in \mathscr{X}^*$ 使得 $\inf\left\langle q^*, \Psi^0(x, y; d^0)\right\rangle = \langle x^*, d^0\rangle\langle q_0^*, p\rangle$ 和 $x^* \in \partial_w^0\psi(x, y)_p$.

证明 由定理 7.2.1、定理 7.1.5 和定理 7.3.2 可以推得本定理. ▯

定理 7.3.4 设 $x \in \mathscr{X}$，$y \in \psi(x)$，$p \in \mathrm{int}K$，$\overline{D} = \mathscr{X}$. 若 $\Psi^0(x, y, \theta)$是K-有界的，$\partial_w^0\psi(x, y)_p \neq \varnothing$，则 $\partial_w^0\psi(x, y)_p$ 是闭凸集.

证明 与定理 7.2.6 的证明类似. ▯

定理 7.3.5 设 $x \in \mathscr{X}$，$y \in \psi(x)$，$p \in \mathrm{int}K$，$\overline{D} = \mathscr{X}$. 若 $\Psi^0(x, y, \theta)$是K-有界的，则 $x^* \in \partial_w^0\psi(x, y)_p$ 当且仅当存在 $q^* \in K^*\backslash\{\theta\}$，使得

$$\langle q^*, y'\rangle \geqslant \langle x^*, d\rangle\langle q^*, p\rangle \quad \forall d \in \mathscr{X}, \forall y' \in \psi_w^0(x, y; d).$$

证明 与定理 7.2.3 的证明类似. ▯

推论 7.3.6 设 $x \in \mathscr{X}$，$y \in \psi(x)$，$\overline{D} = \mathscr{X}$，并且 p^1，$p^2 \in \mathrm{int}K$. 若$\Psi^0(x, y, \theta)$是K-有界的，$x^* \in \partial_w^0\psi(x, y)_{p^1}$，则存在 $q^* \in K^*\backslash\{\theta\}$ 使得 $\alpha x^* \in \partial_w^0\psi(x, y)_{p^2}$，其中 $\alpha = \langle q^*, p^1\rangle/\langle q^*, p^2\rangle$.

证明 由定理 7.3.5 可得证. ▯

定理 7.3.7 设 $x \in \mathscr{X}$，$y \in \psi(x)$，$\overline{D} = \mathscr{X}$，$p \in \mathrm{int}K$. 若 $\Psi^0(x, y, \theta) \subset K$，$\mathrm{int}(K\text{-epi}\Psi^0) \neq \varnothing$，$\partial_w^0\psi(x, y)_p \neq \varnothing$，则对任给的 $d \in \mathscr{X}$ 存在一 $q^* \in K^*\backslash\{\theta\}$，使得

$$\begin{aligned}&\mathrm{int}\left\langle q^*, \psi_w^0(x, y; d)\right\rangle \\ &= \max\left\{\langle x^*, d\rangle\langle q^0, p\rangle \,|\, x^* \in \partial_w^0\psi(x, y)_p\right\}.\end{aligned}$$

证明 与定理 7.2.6 的证明类似. ▯

定义 7.3.4 设 $\psi_1(x)$和 $\psi_2(x)$($\forall x \in \mathscr{X}$)是 $\mathscr{Y}$ 中的非空集合，$\psi_1: \mathscr{X} \to 2^{\mathscr{Y}}$，$x \mapsto \psi_1(x)$ 和 $\psi_2: \mathscr{X} \to 2^{\mathscr{Y}}$，$x \mapsto \psi_2(x)$ 是集值映射，定义集值映射 $\Theta: \mathscr{X} \times \mathscr{Y} \to 2^{\mathscr{Y}} \times 2^{\mathscr{Y}}$，$(x, y) \mapsto \Theta(x, y)$，

$$\Theta(x, y) = \{(y^1, y^2) \in \mathscr{Y} \times \mathscr{Y} \,|\, y^1 \in \psi_1(x),$$

$$y^2 \in \psi_2(x),\ y = y^1 + y^2\}. \qquad (7.3.1)$$

若点列 $\{(x^k, y^k)\}$ 收敛于点 (x^0, y^0)，则对任何 $(y_1^0, y_2^0) \in \Theta(x^0, y^0)$，存在一个序列 $\{(y_k^1, y_k^2)\} \subset \Theta(x^k, y^k)$ 收敛于点 $(y_1^0, y_2^0) \in \Theta(x^0, y^0)$，则称 $\Theta(x, y)$ 在点 (x^0, y^0) 处是局部有界的.

引理 7.3.8　设 $\psi_1(x)$ 和 $\psi_2(x)$ $(\forall x \in \mathscr{X})$ 是 $\mathscr{Y}$ 中的非空集合，$\psi_1: \mathscr{X} \to 2^{\mathscr{Y}}$，$x \mapsto \psi_1(x)$ 和 $\psi_2: \mathscr{X} \to 2^{\mathscr{Y}}$，$x \mapsto \psi_2(x)$ 是集值映射. 又设 $x \in \mathscr{X}$，$y \in \psi_1(x) + \psi_2(x)$. 若由(7.3.1)定义的集值映射 Θ 在点 (x, y) 处局部有界，则

$$\bigcup_{(y^1, y^2) \in \Theta(x, y)} \left[(\psi_1)_w^0(x, y^1; d) + (\psi_2)_w^0(x, y^2; d)\right]$$

$$\subset (\psi_1 + \psi_2)_w^0(x, y; d), \qquad (7.3.2)$$

其中 $d \in \{d \in \mathscr{X} \mid (\psi_1 + \psi_2)_w^0(x, y; d) \neq \varnothing\}$.

证明　考虑任何收敛于 $(x, y, 0)$ 的序列

$$\{(x^k, y^k, t_k)\} \subset K\text{-epi}(\psi_1 + \psi_2) \times (0, +\infty)\ (k = 1, 2, \cdots).$$

对任何 $(y^1, y^2) \in \Theta(x, y)$，设 $r^1 \in \Psi_1^0(x, y^1, d)$，$r^2 \in \Psi_2^0(x, y^2, d)$，即 $(d, r^1) \in G_{K\text{-epi}\psi_1}(x, y^1)$，$(d, r^1) \in G_{K\text{-epi}\psi_2}(x, y^2)$，由定义 7.3.3 可知，存在序列 $\{(y_k^1, y_k^2)\} \subset \Theta(x^k, y^k)$ 收敛于 (y^1, y^2) $(k \to \infty)$，其中 $y^k = y_k^1 + y_k^2$，$y = y^1 + y^2$. 因此，由 G-切锥的定义知存在序列 $\{r_k^1\}$ 和 $\{r_k^2\}$，$r_k^1 \to r^1$，$r_k^2 \to r^2$，使得

$$y_k^1 + t_k r_k^1 \in \psi_1(x^k + t_k d) + K,$$

$$y_k^2 + t_k y_k^2 \in \psi_2(x^k + t_k d) + K,$$

从而得

$$y_k^1 + y_k^2 + t_k r_k^1 + t_k r_k^2 = y^k + t_k(r_k^1 + r_k^2) \in \psi_1(x^k + t_k d) + K$$

$$+ \psi_2(x^k + t_k d) + K \subset (\psi_1 + \psi_2)(x^k + t_k d) + K.$$

因为 $r_k^1 + r_k^2 \to r^1 + r^2$，所以 $(d, r^1 + r^2) \in G_{K\text{-epi}(\psi_1 + \psi_2)}(x, y)$，

故(7.3.2)成立. ▯

定理 7.3.9　设 $\boldsymbol{x} \in \mathscr{X}$, $\boldsymbol{y} \in \psi(\boldsymbol{x})$, $\overline{D} = \mathscr{X}$, $\boldsymbol{p} \in \operatorname{int}K$. 若 $\Psi^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0})$是$K$-有界的,则 $\psi_w^0(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d})$关于 $\boldsymbol{d}$ 在 $\mathscr{X}$ 上是K-凸的.

证明　因为$\Psi^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0})$是$K$-有界的,可以推出对任何 $\boldsymbol{d} \in \mathscr{X}$, $\Psi^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d})$ 是 K-有界的. 由定理 7.3.1 可知 $\Psi^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d})$ $(\forall \boldsymbol{d} \in \mathscr{X})$ 是 K-闭的,由引理 6.5.1 知$\Psi^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d})$是 K-半紧的,再由引理 6.5.2 便得到

$$\Psi^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}) \subset \psi_w^0(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}) + K \quad \forall \boldsymbol{d} \in \mathscr{X}. \tag{7.3.3}$$

又对任意的 $\boldsymbol{d}^1$, $\boldsymbol{d}^2 \in \mathscr{X}$ 和 $\lambda \in (0, 1)$, 由定理 7.3.1 有

$$\begin{aligned} &\lambda \Psi^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}^1) + (1-\lambda) \Psi^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}^2) \\ &\quad \in \Psi^0\left(\boldsymbol{x}, \boldsymbol{y}, \lambda \boldsymbol{d}^1 + (1-\lambda)\boldsymbol{d}^2\right). \end{aligned} \tag{7.3.4}$$

从(7.3.3)和(7.3.4)则得

$$\begin{aligned} &\lambda\psi_w^0(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}^1) + (1-\lambda)\psi_w^0(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}^2) \\ &\subset \lambda\Psi^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}^1) + (1-\lambda) \Psi^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}^2) \\ &\subset \Psi^0(\boldsymbol{x}, \boldsymbol{y}, \lambda\boldsymbol{d}^1 + (1-\lambda)\boldsymbol{d}^2) \\ &\subset \psi_w^0\left(\boldsymbol{x}, \boldsymbol{y}; \lambda\boldsymbol{d}^1 + (1-\lambda)\boldsymbol{d}^2\right) + K. \end{aligned}$$ ▯

定理 7.3.10　设 $\psi_1(\boldsymbol{x})$和 $\psi_2(\boldsymbol{x})$ $(\forall \boldsymbol{x} \in \mathscr{X})$是 $\mathscr{Y}$ 中的非空集合, $\psi_1: \mathscr{X} \to 2^{\mathscr{Y}}$, $\boldsymbol{x} \mapsto \psi_1(\boldsymbol{x})$和 $\psi_2: \mathscr{X} \to 2^{\mathscr{Y}}$, $\boldsymbol{x} \mapsto \psi_2(\boldsymbol{x})$是集值映射. 又设 $\boldsymbol{x} \in \mathscr{X}$, $\boldsymbol{y} \in \psi_1(\boldsymbol{x}) + \psi_2(\boldsymbol{x})$, $\Psi_1^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0})$ 和$\Psi_2^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0})$是 K-有界的, $\overline{D}_1 = \overline{D}_2 = \mathscr{X}$ $\left(\overline{D}_1 = \left\{\boldsymbol{d} \mid (\psi_1)_w^0(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}) \neq \varnothing\right\}, \overline{D}_2 = \left\{\boldsymbol{d} \mid (\psi_2)_w^0(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}) \neq \varnothing\right\}\right)$, $\operatorname{int}(K\text{-epi}\Psi_1^0) \neq \varnothing$ 或者 $\operatorname{int}(K\text{-epi}\Psi_2^0) \neq \varnothing$. 若按(7.3.1)定义的 $\Theta(\boldsymbol{x}, \boldsymbol{y})$在点$(\boldsymbol{x}, \boldsymbol{y})$处是局部有界的,则

$$\partial_w^0(\psi_1 + \psi_2)(\boldsymbol{x}, \boldsymbol{y})_{\boldsymbol{p}}$$

$$\subset \bigcup_{(y^1, y^2)\in\Theta(x, y)} \left(\partial_w^0\psi_1(x, y^1)_p + \partial_w^0\psi_2(x, y^2)_p\right).$$

证明 取 $x^* \in \partial_w^0(\psi_1+\psi_2)(x, y)_p$，不妨设 $\text{int}(K\text{-epi}\Psi_2^0) \neq \varnothing$. 对任意的 $(y^1, y^2) \in \Theta(x, y)$，作集合

$$S_1 = \{(d, r) \in \mathscr{X} \times \mathscr{Y} \mid d \in \mathscr{X}, r^1 \in \Psi_1^0(x, y^1, d),$$

$$q^1 \in K, r = -r^1 - q^1 + \langle x^*, d\rangle p\},$$

$$S_2 = \{(d, r) \in \mathscr{X} \times \mathscr{Y} \mid d \in \mathscr{X}, r^2 \in \Psi_2^0(x, y^2, d),$$

$$q^2 \in K, r = r^2 + q^2\}.$$

由定理 7.3.1 可知 S_1 和 S_2 是非空凸集，并且 $\text{int}S_2 \neq \varnothing$. 以下证明 $S_1 \cap \text{int}S_2 = \varnothing$. 事实上，反之，假若存在 $(\bar{d}, \bar{r}) \in S_1 \cap \text{int}S_2$，我们有

$$\bar{r} = -r^1 - q^1 + \langle x^*, \bar{d}\rangle p,$$

并且存在 $\alpha > 0$，使 $(\bar{d}, \bar{r} - \alpha p) \in S_2$. 因此，有 $r^2 \in \Psi_2^0(x, y, \bar{d})$ 使得 $\bar{r} - \alpha p = r^2 + q^2$，其中 $q^2 \in K$，从而

$$-r^1 - q^1 + \langle x^*, \bar{d}\rangle p = r^2 + q^2 + \alpha p,$$

即

$$-\left(r^1 + r^2 - \langle x^*, \bar{d}\rangle p\right) = q^1 + q^2 + \alpha p \in \text{int}K.$$

据此，由引理 7.3.8 得 $r^1 + r^2 \in (\Psi_1 + \Psi_2)^0(x, y, \bar{d})$. 从定理 7.3.2 知 $x^* \notin \partial_w^0(\psi_1 + \psi_2)(x, y)_p$，这与已设矛盾. 现在，利用定理 1.3.7 可知存在 $(\bar{x}^*, \bar{y}^*) \in \mathscr{X}^* \times \mathscr{Y}^*$ 和 $(\bar{x}^*, \bar{y}^*) \neq \boldsymbol{0}$，使

$$\langle \bar{x}^*, d^{S_1}\rangle + \langle \bar{y}^*, r^{S_1}\rangle \geqslant \beta \geqslant \langle \bar{x}^*, d^{S_2}\rangle + \langle \bar{y}^*, r^{S_2}\rangle$$

$$\forall (d^{S_1}, r^{S_1}) \in S_1, \forall (d^{S_2}, r^{S_2}) \in S_2. \tag{7.3.5}$$

因为 $(\boldsymbol{0}, \boldsymbol{0}) \in S_1 \cap S_2$，所以从(7.3.5)得 $\beta = 0$. 假设 $\bar{y}^* = \boldsymbol{0}$，则 $\bar{x}^* \neq \boldsymbol{0}$，从(7.3.5)有

$$\langle \bar{x}^*, d\rangle \geqslant 0 \quad \forall d \in \mathscr{X},$$

从而得 $\bar{x}^* = 0$，导致矛盾. 因此，有 $\bar{y}^* \neq 0$. 在(7.3.5)中设 $d^{S_2} = 0$, $r^{S_2} = q \in K$，得

$$\langle -\bar{y}^*, q\rangle \geqslant 0 \quad \forall q \in K,$$

从而 $-\bar{y}^* \in K^* \backslash \{0\}$. 于是，由定理 1.4.11 得 $\langle -\bar{y}^*, p\rangle > 0$. 在(7.3.5)中令 $q^* = -\bar{y}^*$, $d^{S_2} = d$, $r^{S_2} = r^2$, $x_2^* = \bar{x}^* / \langle q^*, p\rangle$，我们有

$$\langle q^*, r^2\rangle \geqslant \langle x_2^*, d\rangle\langle q^*, p\rangle \quad \forall d \in \mathscr{X}, \forall r^2 \in \Psi_2^0(x, y_2, d).$$

由此，根据定理 7.3.5 得到 $x_1^* \in \partial_w^0 \psi_2(x, y_1)_p$. 在(7.3.5)中再令 $q^* = -\bar{y}^*$, $d^{S_1} = d$, $r^{S_1} = -r^1 + \langle x^*, d\rangle p$, $x_1^* = x^* - x_2^*$，则有

$$\langle q^*, r^1\rangle \geqslant \langle x_1^*, d\rangle\langle q^*, p\rangle, \quad \forall d \in \mathscr{X},$$

$$\forall r^1 \in \Psi_1^0(x, y_1, d).$$

于是，由定理 7.3.5 又得到 $x_1^* \in \partial_w^0 \psi_1(x, y_1)_p$. ▯

定理 7.3.11 设 $x \in \mathscr{X}$, $y \in \psi(x)$, $p \in \operatorname{int} K$, $\overline{D} = \mathscr{X}$，并且 $\Psi^0(x, y, 0)$ 是 K-有界的. 若 (x, y) 是 ψ 的局部弱有效解，则 $0 \in \partial_w^0 \psi(x, y)_p$.

证明 与定理 7.2.9 的证明类似. ▯

§7.4 算子锥次微分

在前面，我们都是通过线性泛函来引进集值映射的锥次微分的. 本节将从线性算子来定义集值映射的锥次微分.

设 $\mathscr{X}$ 和 $\mathscr{Y}$ 是局部凸线性拓扑空间，$L(\mathscr{X}, \mathscr{Y})$ 表示 $\mathscr{X}$ 到 $\mathscr{Y}$ 的所有连续线性算子组成的集合，$\psi: \mathscr{X} \to 2^{\mathscr{Y}}$, $x \mapsto \psi(x)$ 是集值映射，$K \subset \mathscr{Y}$ 是内部非空的尖闭凸锥.

定义 7.4.1 设集合 $S \subset \mathscr{X}$ 非空，集值映射 ψ 在 S 上是 K

凸的，又点 $x^0 \in S$，$y^0 \in \psi(x^0)$.

(1) 若存在 $\mathscr{A} \in L(\mathscr{X}, \mathscr{Y})$ 使得

$$y^0 - \mathscr{A}(x^0) \in \mathscr{E}\left(\bigcup_{x \in S}\left[\psi(x) - \mathscr{A}(x)\right], K\right),$$

则称 $\mathscr{A}$ 是集值映射 ψ 在点 (x^0, y^0) 处的算子 K-次梯度. ψ 在点 (x^0, y^0) 处的所有算子 K-次梯度组成的集合称为 ψ 在点 (x^0, y^0) 处的算子 K-次微分，记作 $\partial\psi(x^0, y^0)_L$.

(2) 若存在 $\mathscr{A} \in L(\mathscr{X}, \mathscr{Y})$ 使得

$$y^0 - \mathscr{A}(x^0) \in \mathscr{E}_w\left(\bigcup_{x \in S}[\psi(x) - \mathscr{A}(x)], K\right),$$

则称 $\mathscr{A}$ 是集值映射 ψ 在点 (x^0, y^0) 处的算子 K-弱次梯度. ψ 在点 (x^0, y^0) 处的所有算子 K-弱次梯度组成的集合称为 ψ 在点 (x^0, y^0) 处的算子 K-弱次微分，记作 $\partial_w\psi(x^0, y^0)_L$.

由定义 6.1.3 和定义 7.4.1 易知有

$$\partial\psi(x^0, y^0)_p \subset \partial\psi(x^0, y^0)_L, \quad \partial_w\psi(x^0, y^0)_p \subset \partial_w\psi(x^0, y^0)_L. \tag{7.4.1}$$

下面是存在性定理.

定理 7.4.1　设 $S \subset \mathscr{X}$ 是内部非空的凸集，$\psi(x) \neq \varnothing (\forall x \in S)$，$x \in \text{int}S$，$p \in \text{int}K$. 又设 ψ 在 S 上是 K-凸的，$\text{int}(K\text{-epi}\psi) \neq \varnothing$，$y \in \mathscr{E}(\psi(x), K)$.

(1) $\partial_w\psi(x, y)_L \neq \varnothing$.

(2) 若 ψ 在 x 处是 K-严格凸的，则 $\partial\psi(x, y)_L \neq \varnothing$.

证明　由定理 6.2.4 和(7.4.1)即可得证. □

定理 7.4.2　设 $S \subset \mathscr{X}$ 是内部非空的凸集，$\psi(x) \neq \varnothing$ $(\forall x \in S)$，$x \in S$，$y \in \psi(x)$. 若 ψ 在 S 是上 K-凸的，则 $\mathscr{A} \in \partial_w\psi(x, y)_L$ 当且仅当存在 $q^* \in K^* \setminus \{0\}$，使得

$$\langle q^*, y' - y\rangle \geqslant \langle q^*, \mathscr{A}(x') - \mathscr{A}(x)\rangle$$

$$\forall x' \in S, \forall y' \in \psi(x).$$

证明 与定理 6.2.11 的(2)和定理 6.2.12 的证明类似.▯

定理 7.4.3 设 $S \subset \mathscr{X}$ 是内部非空的凸集，$\psi(\boldsymbol{x}) \neq \varnothing$ ($\forall \boldsymbol{x} \in S$)，$\boldsymbol{x} \in S$，$\boldsymbol{y} \in \psi(\boldsymbol{x})$. 若 ψ 在 S 上是 K-凸的，$\partial_w \psi(\boldsymbol{x}, \boldsymbol{y})_L \neq \varnothing$，则 $\partial_w \psi(\boldsymbol{x}, \boldsymbol{y})_L$ 是闭凸集.

证明 与定理 6.3.1 和定理 6.3.2 的(1)的证明类似.▯

定理 7.4.4 设 $S \subset \mathscr{X}$ 是内部非空的凸集，$\boldsymbol{x} \in S$，$\psi_1(\boldsymbol{x})$和 $\psi_2(\boldsymbol{x})$ ($\forall \boldsymbol{x} \in S$) 非空，

$$(\psi_1 + \psi_2)(\boldsymbol{x}) = \{\boldsymbol{y}^1 + \boldsymbol{y}^2 \mid \boldsymbol{y}^1 \in \psi_1(\boldsymbol{x}),\ \boldsymbol{y}^2 \in \psi_2(\boldsymbol{x}),\ \boldsymbol{x} \in S\},$$

$$\boldsymbol{y}^1 \in \psi_1(\boldsymbol{x}),\ \boldsymbol{y}^2 \in \psi_2(\boldsymbol{x}).$$

$$\psi_1 : \mathscr{X} \to 2^{\mathscr{Y}},\ \boldsymbol{x} \mapsto \psi_1(\boldsymbol{x}),\ \psi_2 : \mathscr{X} \to 2^{\mathscr{Y}},\ \boldsymbol{x} \mapsto \psi_2(\boldsymbol{x}),$$

$\psi_1 + \psi_2 : \mathscr{X} \to 2^{\mathscr{Y}}$, $\boldsymbol{x} \mapsto (\psi_1 + \psi_2)(\boldsymbol{x})$ 是集值映射.
若 ψ_1 和 ψ_2 在 S 上是 K-凸的，ψ_1 或 ψ_2 在 $\boldsymbol{x}$ 处是连通的，则

$$\partial_w(\psi_1 + \psi_2)(\boldsymbol{x}, \boldsymbol{y}^1 + \boldsymbol{y}^2)_L \subset \partial_w \psi_1(\boldsymbol{x}, \boldsymbol{y}^1)_L + \partial_w \psi_2(\boldsymbol{x}, \boldsymbol{y}^2)_L.$$

证明 与定理 6.4.2 的证明类似.▯

定理 7.4.5 设 $S \subset \mathscr{X}$ 是非空凸集，$\boldsymbol{x} \in S$，$\boldsymbol{y} \in \psi(\boldsymbol{x})$，$\Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0})$ 和 $\psi'(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d})$按定义 6.5.1 所确定. 若 ψ 在 S 上是 K-凸的，$\hat{D} = \mathscr{X}$，$\Psi(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0})$ 是 K-有界的，则

$$\partial \psi(\boldsymbol{x}, \boldsymbol{y})_L$$

$$\supset \left\{\mathscr{A} \in L(\mathscr{X}, \mathscr{Y}) \,\middle|\, \boldsymbol{0} \in \mathscr{E}\left(\bigcup_{\boldsymbol{d} \in \mathscr{X}} [\psi'(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}) - \mathscr{A}(\boldsymbol{d})], K\right)\right\},$$

$$\partial_w \psi(\boldsymbol{x}, \boldsymbol{y})_L$$

$$= \left\{\mathscr{A} \in L(\mathscr{X}, \mathscr{Y}) \,\middle|\, \boldsymbol{0} \in \mathscr{E}_w\left(\bigcup_{\boldsymbol{d} \in \mathscr{X}} [\psi'(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}) - \mathscr{A}(\boldsymbol{d})], K\right)\right\}.$$

证明 与定理 6.5.3 的证明类似.▯

通过 C-切锥和 G-切锥，也可以由算子构成集值映射的锥次微分.

定义 7.4.2　设点 $x^0 \in \mathscr{X}$, $y^0 \in \psi(x^0)$, $d \in \mathscr{X}$, $\psi_w^C(x^0, y^0; d)$是 ψ 在点(x^0, y^0)处沿方向 d 的 CK-弱方向导数,则称集合

$$\partial_w^C\psi(x^0, y^0)_L = \left\{\mathscr{A} \in L(\mathscr{X}, \mathscr{Y}) \;\middle|\; 0 \in \mathscr{E}_w\left(\bigcup_{d \in \mathscr{X}}\left[\psi_w^C(x^0, y^0; d) - \mathscr{A}(d)\right], K\right)\right\}$$

是集值映射 ψ 在点(x^0, y^0)处的算子 CK-弱次微分,称 $\mathscr{A} \in \partial_w^C\psi(x^0, y^0)_L$ 是 ψ 在点(x^0, y^0)处的算子 CK-弱次梯度. 若 $\partial_w^C\psi(x^0, y^0)_L \neq \varnothing$,则称 ψ 在点(x^0, y^0)处是算子 CK-弱次可微的. 记

$$\partial_w^C\psi(x^0)_L = \{\partial_w^C(x^0, y^0)_L \mid y^0 \in \psi(x^0)\},$$

并称它是 ψ 在点 x^0 处的算子 CK-弱次微分.

定理 7.4.6　设 $x \in \mathscr{X}$, $y \in \psi(x)$, $\hat{D} = \mathscr{X}$. 若$\Psi(x, y, 0)$是 K-有界的,则

$$\begin{aligned}\partial_w^C\psi(x, y)_L &= \left\{\mathscr{A} \in L(\mathscr{X}, \mathscr{Y}) \;\middle|\; 0 \in \mathscr{E}_w\left(\bigcup_{d \in \mathscr{X}}\left[\Psi(x, y, d) - \mathscr{A}(d)\right], K\right)\right\} \\ &= \left\{\mathscr{A} \in L(\mathscr{X}, \mathscr{Y}) \;\middle|\; 0 \in \mathscr{E}_w\left(\bigcup_{(d, r) \in T^C_{K\text{-epi}\psi}(x, y)}\left[r - \mathscr{A}(d)\right], K\right)\right\}.\end{aligned}$$

证明　与定理 7.2.1 的证明类似. □

定理 7.4.7　设 $x \in \mathscr{X}$, $y \in \psi(x)$, $p \in \mathrm{int}K$, $\hat{D} = \mathscr{X}$. 又设 $\Psi(x, y, 0) \subset K$, $\mathrm{int}(K\text{-epi}\Psi) \neq \varnothing$. 若存在 $\bar{d} \in \mathscr{X}$ 和 $\bar{q}^* \in K^* \setminus \{0\}$,使得 $\inf\langle \bar{q}^*, \Psi(x, y, \bar{d})\rangle > -\infty$,则 $\partial_w^C\psi(x, y)_L \neq \varnothing$.

证明　由定理 7.2.2 知存在 $x^* \in \partial_w^C\psi(x, y)_p$. 作线性算子

$$\mathscr{A}: \mathscr{X} \to \mathscr{Y},\ \mathscr{A}(x) = \langle x^*, x\rangle p \quad \forall x \in S,$$

则 $\mathscr{A} \in \partial_w^C\psi(x, y)_L$,故得证. □

定理 7.4.8　设 $x \in \mathscr{X}$, $y \in \psi(x)$, $\hat{D} = \mathscr{X}$. 又设$\Psi(x, y,$

$\boldsymbol{0}$)是 K-有界的.

(1) $\mathscr{A} \in \partial_w^C \phi(\boldsymbol{x}, \boldsymbol{y})_L$ 当且仅当存在 $\boldsymbol{q}^* \in K^* \setminus \{\boldsymbol{0}\}$,使得

$$\langle \boldsymbol{q}^*, \boldsymbol{y}' \rangle \geqslant \langle \boldsymbol{q}^*, \mathscr{A}(\boldsymbol{d}) \rangle \quad \forall \boldsymbol{d} \in \mathscr{X}, \forall \boldsymbol{y}' \in \phi_w^C(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}).$$

(2) 若 $\partial_w^C \phi(\boldsymbol{x}, \boldsymbol{y})_L \neq \varnothing$,则 $\partial_w^C \phi(\boldsymbol{x}, \boldsymbol{y})_L$ 是闭凸集.

(3) 若($\boldsymbol{x}, \boldsymbol{y}$)是 ϕ 的局部弱有效解,则 $\boldsymbol{0} \in \partial_w^C \phi(\boldsymbol{x}, \boldsymbol{y})_L$.

证明 (1) 与定理 7.2.3 的证明类似.

(2) 与定理 7.2.5 的证明类似.

(3) 与定理 7.2.9 的证明类似. □

定义 7.4.3 设点 $\boldsymbol{x}^0 \in \mathscr{X}$, $\boldsymbol{y}^0 \in \phi(\boldsymbol{x}^0)$, $\boldsymbol{d} \in \mathscr{X}$. $\phi_w^0(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d})$ 是集值映射 ϕ 在点($\boldsymbol{x}^0, \boldsymbol{y}^0$)处沿 $\boldsymbol{d}$ 的 GK-弱方向导数,则称集合

$$\partial_w^0 \phi(\boldsymbol{x}^0, \boldsymbol{y}^0)_L$$
$$= \left\{ \mathscr{A} \in L(\mathscr{X}, \mathscr{Y}) \,\middle|\, \boldsymbol{0} \in \mathscr{E}_w\left(\bigcup_{\boldsymbol{d} \in \mathscr{X}} \left[\phi_w^0(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d}) - \mathscr{A}(\boldsymbol{d}) \right], K \right) \right\}$$

是集值映射 ϕ 在点($\boldsymbol{x}^0, \boldsymbol{y}^0$)处的算子 GK-弱次微分,称 $\mathscr{A} \in \partial_w^0 \phi(\boldsymbol{x}^0, \boldsymbol{y}^0)$ 是 ϕ 在点($\boldsymbol{x}^0, \boldsymbol{y}^0$)处的算子 GK-弱次梯度.若 $\partial_w^0 \phi(\boldsymbol{x}^0, \boldsymbol{y}^0)_L \neq \varnothing$,则称 ϕ 在点($\boldsymbol{x}^0, \boldsymbol{y}^0$)处是算子 GK-弱次可微的.记 $\partial_w^0 \phi(\boldsymbol{x}^0)_L = \{\partial_w^C \phi(\boldsymbol{x}^0, \boldsymbol{y}^0)_L \mid \boldsymbol{y}^0 \in \phi(\boldsymbol{x}^0)\}$,并称它是 ϕ 在点 $\boldsymbol{x}^0$ 处的算子 GK-弱次微分.

定理 7.4.9 设 $\boldsymbol{x} \in \mathscr{X}$, $\boldsymbol{y} \in \phi(\boldsymbol{x})$, $\overline{D} = \mathscr{X}$. 若 $\Psi^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0})$ 是 K-有界的,则

$$\partial_w^0 \phi(\boldsymbol{x}, \boldsymbol{y})_L$$
$$= \left\{ \mathscr{A} \in L(\mathscr{X}, \mathscr{Y}) \,\middle|\, \boldsymbol{0} \in \mathscr{E}_w\left(\bigcup_{\boldsymbol{d} \in \mathscr{X}} \left[\Psi^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{d}) - \mathscr{A}(\boldsymbol{d}) \right], K \right) \right\}$$
$$= \left\{ \mathscr{A} \in L(\mathscr{X}, \mathscr{Y}) \,\middle|\, \boldsymbol{0} \in \mathscr{E}_w\left(\bigcup_{(\boldsymbol{d}, \boldsymbol{r}) \in G_{K\text{-epi}\phi}(\boldsymbol{x}, \boldsymbol{y})} \left[\boldsymbol{r} - \mathscr{A}(\boldsymbol{d}) \right], K \right) \right\}.$$

证明 与定理 7.3.2 的证明类似. □

定理 7.4.10 设 $\boldsymbol{x} \in \mathscr{X}$, $\boldsymbol{y} \in \phi(\boldsymbol{x})$, $\boldsymbol{p} \in \operatorname{int}K$, $\overline{D} = \mathscr{X}$. 又

$\Psi^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0}) \subset K$, $\mathrm{int}(K\text{-epi}\Psi^0) \neq \varnothing$. 若存在 $\overline{\boldsymbol{d}} \in \mathscr{X}$ 和 $\overline{\boldsymbol{q}}^* \in K^* \setminus \{\boldsymbol{0}\}$, 使得 $\inf\langle \overline{\boldsymbol{q}}^*, \psi_w^0(\boldsymbol{x}, \boldsymbol{y}; \overline{\boldsymbol{d}})\rangle > -\infty$, 则 $\partial_w^0\psi(\boldsymbol{x}, \boldsymbol{y})_L \neq \varnothing$.

证明　与定理 7.3.3 的证明类似. □

定理 7.4.11　设 $\boldsymbol{x} \in \mathscr{X}$, $\boldsymbol{y} \in \psi(\boldsymbol{x})$, $\overline{D} = \mathscr{X}$. 又设 $\Psi^0(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{0})$ 是 K-有界的.

(1) 若 $\partial_w^0\psi(\boldsymbol{x}, \boldsymbol{y})_L \neq \varnothing$, 则 $\partial_w^0\psi(\boldsymbol{x}, \boldsymbol{y})_L$ 是闭凸集.

(2) $\mathscr{A} \in \partial_w^0\psi(\boldsymbol{x}, \boldsymbol{y})_L$ 当且仅当存在 $\boldsymbol{q}^* \in K^* \setminus \{\boldsymbol{0}\}$, 使得

$$\langle \boldsymbol{q}^*, \boldsymbol{y}'\rangle \geqslant \langle \boldsymbol{q}^*, \mathscr{A}(\boldsymbol{d})\rangle \quad \forall \boldsymbol{d} \in \mathscr{X}, \forall \boldsymbol{y}' \in \psi_w^0(\boldsymbol{x}, \boldsymbol{y}; \boldsymbol{d}).$$

(3) 若 $(\boldsymbol{x}, \boldsymbol{y})$ 是 ψ 的局部弱有效解, 则 $\boldsymbol{0} \in \partial_w^0\psi(\boldsymbol{x}, \boldsymbol{y})_L$.

证明　(1), (2)和(3)依次与定理 7.3.4, 定理 7.3.5 和定理 7.3.11 的证明类似. □

第 8 章　集值映射的切导数和切微分

这一章将进一步研究集值映射在其他意义下的导数和微分问题. 与前面介绍的次微分不同的是, 本章将借助相应图象的切锥或法锥, 来定义集值映射的导数和微分. 在本章的最后部分, 我们特别讨论有限维空间中集值映射的广义微分问题.

§8.1　切导数和切微分

设 $\mathscr{B}$ 和 $\mathscr{D}$ 是 Banach 空间, $\mathscr{B}^*$ 和 $\mathscr{D}^*$ 分别是 $\mathscr{B}$ 和 $\mathscr{D}$ 的对偶空间, $\psi: \mathscr{B} \to 2^{\mathscr{D}}$, $x \mapsto \psi(x)$ 是集值映射(对于 $\boldsymbol{x} \in \mathscr{B}$, 集合 $\psi(x) \neq \varnothing$, 且至少有一点 $\boldsymbol{x} \in \mathscr{B}$ 对应的 $\psi(\boldsymbol{x})$ 是非单点集, 除非特别申明 ψ 是单值映射). 记 ψ 的有效域为

$$\operatorname{dom} \psi = \{\boldsymbol{x} \in \mathscr{B} \mid \psi(\boldsymbol{x}) \neq \varnothing\},$$

ψ 的图象为

$$\operatorname{graph} \psi = \{(\boldsymbol{x}, \boldsymbol{y}) \mid \boldsymbol{y} \in \psi(\boldsymbol{x})\},$$

ψ 的象为

$$\operatorname{Im} \psi = \bigcup_{x \in \mathscr{B}} \psi(\boldsymbol{x}) = \bigcup_{x \in \operatorname{dom}\psi} \psi(\boldsymbol{x}).$$

又记 ψ 的逆集值映射为: $\psi^{-1}: \mathscr{D} \to 2^{\mathscr{B}}$, $\boldsymbol{y} \mapsto \psi^{-1}(\boldsymbol{y})$. $\boldsymbol{x} \in \psi^{-1}(\boldsymbol{y})$ 当且仅当 $\boldsymbol{y} \in \psi(\boldsymbol{x})$.

显然有下面关系:

$$\operatorname{dom} \psi^{-1} = \operatorname{Im} \psi, \ \operatorname{Im} \psi^{-1} = \operatorname{dom} \psi.$$

这一节均设 graph ψ 是 $\mathscr{B}\times\mathscr{D}$ 中的凸集，点 $(x, y)\in$ graph ψ，即 $y=\psi(x)$. 显然，集合 graph ψ 在点 (x, y) 处的切锥 $T_{\text{graph}\psi}(x, y)$ 和法锥 $N_{\text{graph}\psi}(x, y)$ 是空间 $\mathscr{B}\times\mathscr{D}$ 中的闭凸集.

定义 8.1.1 设点 $x^0\in\mathscr{B}$，$y^0\in\psi(x^0)$，考虑集值映射 $\psi^T(x^0, y^0): \mathscr{B}\to 2^{\mathscr{D}}$，$u\mapsto\psi^T(x^0, y^0)(u)$. 若对任意的 $v\in\psi^T(x^0, y^0)(u)$ 当且仅当

$$(u, v)\in T_{\text{graph}\psi}(x^0, y^0),$$

则称 $\psi^T(x^0, y^0)$ 是 ψ 在点 (x^0, y^0) 处的切导数. 考虑集值映射 $\partial^T\psi(x^0, y^0): \mathscr{D}^*\to 2^{\mathscr{B}^*}$，$q^*\mapsto\partial^T\psi(x^0, y^0)(q^*)$. 若对任意的 $p^*\in\partial^T\psi(x^0, y^0)(q^*)$ 当且仅当

$$(p^*, -q^*)\in N_{\text{graph}\psi}(x^0, y^0),$$

则称 $\partial^T\psi(x^0, y^0)$ 是 ψ 在点 (x^0, y^0) 处的切微分.

由定义 3.2.2 和定义 8.1.1 易得以下结果.

定理 8.1.1 设 $(x, y)\in$ graph ψ，则 $p^*\in\partial^T\psi(x, y)(q^*)$ 当且仅当

$$\langle q^*, y-y'\rangle\leqslant\langle p^*, x-x'\rangle\quad\forall x'\in\mathscr{B},\quad\forall y'\in\psi(x'),$$

或

$$\langle p^*, u\rangle\leqslant\langle q^*, v\rangle\quad\forall u\in\mathscr{B},\quad\forall v\in\psi^T(x, y)(u).$$

定理 8.1.2 设 $x\in\mathscr{B}$，$y\in\psi(x)$.

(1) $(\psi^{-1})^T(y, x)=\psi^T(x, y)^{-1}$.

(2) $\partial^T(\psi^{-1})(y, x)(p^*)=-(\partial^T\psi(x, y))^{-1}(-p^*)$.

定理 8.1.3 设 $x^1, x^2\in\mathscr{B}$，$y^1\in\psi(x^1)$，$y^2\in\psi(x^2)$，$p_1^*\in\partial^T\psi(x^1, y^1)(q_1^*)$，$p_2^*\in\partial^T\psi(x^2, y^2)(q_2^*)$，则 $\langle q_1^*-q_2^*, y^1-y^2\rangle\leqslant\langle p_1^*-p_2^*, x^1-x^2\rangle$.

定理 8.1.4 设 $x\in\mathscr{B}$，$y\in\psi(x)$.

(1) dom $(\psi^T(x, y))\subset T_{\text{dom}\psi}(x)$.

(2) $\text{Im}(\psi^T(x, y))\subset T_{\text{Im}\psi}(y)$.

设 $K\subset\mathscr{B}$ 是内部非空的尖闭凸锥.

定理 8.1.5 设 $x\in\mathscr{B}$, $y\in\psi(x)$, $p\in\operatorname{int}K$, $d\in\mathscr{B}$, 又 $\Psi(x,y,d)$由(6.5.2)确定，$\widehat{D}=\mathscr{B}$, $\Psi(x,y,0)$是K-有界的.

(1) $\psi^T(x,y)(d)\subset\Psi(x,y,d)$.

(2) 若 $x^*\in\partial^T\psi(x,y)(q^*)$, $q^*\in K^*\backslash\{0\}$, K^*是K的对偶锥，则

$$x^*(\langle q^*,p\rangle)^{-1}\in\partial_w\psi(x,y)_p.$$

(3) 若 $x^*\in\partial_w\psi(x,y)_p$, 则存在 $q^*\in K^*\backslash\{0\}$ 使得

$$\langle q^*,p\rangle x^*\in\partial^T\psi(x,y)(q^*).$$

证明 (1) 由定义 8.1.1 和定义 6.5.1 直接得证.

(2) 由定理 8.1.1 和定理 6.2.11 可以推得证明.

(3) 由定理 8.1.1 和定理 6.2.12 可以推证. ▯

定理 8.1.6 设 $x\in\mathscr{B}$, $y\in\psi(x)$, 则 $v\in\psi^T(x,y)(u)$ 当且仅当

$$\lim_{u'\to u}\inf_{t>0}d\left(v,\frac{\psi(x+tu')-y}{t}\right)=0. \tag{8.1.1}$$

证明 因为 $(u,v)\in T_{\text{graph}\psi}(x,y)$ 当且仅当对任何 $\varepsilon_1>0$, $\varepsilon_2>0$, 存在 $u_{\varepsilon_1}\in\mathscr{B}$, $v_{\varepsilon_2}\in\mathscr{D}$, 满足 $\|u_{\varepsilon_1}\|\leqslant\varepsilon_1$, $\|v_{\varepsilon_2}\|\leqslant\varepsilon_2$ 和 $t_0>0$, 使得

$$\left(x+t(u+u_{\varepsilon_1}),\ y+t(v+v_{\varepsilon_2})\right)\in\operatorname{graph}\psi\quad\forall t\in[0,t_0),$$

所以

$$v\in\frac{\psi(x+t(u+u_{\varepsilon_1}))-y}{t}-v_{\varepsilon_2}.$$

由此，我们有

$$\inf_{0<t<t_0}d\left(v,\frac{\psi(x+t(u+u_{\varepsilon_1})-y}{t}\right)\leqslant\varepsilon_2. \tag{8.1.2}$$

由于 graph ψ 是凸集，对任何 $t_1\leqslant t_2$, 有

$$\frac{\psi(x' + t_2 u') - y}{t_2} \subset \frac{\psi(x' + t_1 u') - y'}{t_1}$$

$$\forall x' \in \mathscr{B},\ y' \in \psi(x'),$$

故

$$\lim_{t \to 0} d\left(v, \frac{\psi(x + tu') - y}{t}\right) = \inf_{t > 0} d\left(v, \frac{\psi(x + tu') - y}{t}\right).$$

由上式和(8.1.2)可得，对 $\|u_{\varepsilon_1}\| \leqslant \varepsilon_1$ 有

$$\inf_{0 < t < t_0} d\left(v, \frac{\psi(x + t(u + u_{\varepsilon_1}) - y}{t}\right)$$

$$= \inf_{t > 0} d\left(v, \frac{\psi(x + t(u + u_{\varepsilon_1}) - y}{t}\right)$$

$$= \lim_{t \to 0^+} d\left(v, \frac{\psi(x + t(u + u_{\varepsilon_1}) - y}{t}\right).$$

于是，从上式得知

$$\inf_{\|u' - u\| \leqslant \varepsilon_1} \inf_{t > 0} d\left(v, \frac{\psi(x + t(u + u_{\varepsilon_1}) - y}{t}\right) \leqslant \varepsilon_2.$$

令 $\varepsilon_1 \to 0$, $\varepsilon_2 \to 0$，即得(8.1.1). ▯

定理 8.1.7 设 $x^0, x \in \operatorname{dom} \psi$，则

$$\psi(x^0) - y \subset \psi^T(x, y)(x^0 - x), \quad \forall y \in \psi(x).$$

证明 由于 ψ 在 $\operatorname{dom} \psi$ 上是凸的，故对任意的 $\lambda \in (0, 1)$ 和 $y^0 \in \psi(x^0)$，有

$$(1 - \lambda) y + \lambda y^0 \subset \psi(x + \lambda(x^0 - x)),$$

即

$$y^0 - y \in \frac{\psi\big(x + \lambda(x^0 - x)\big) - y}{\lambda}.$$

由定义 8.1.1 得到 $y^0 - y \in \psi^T(x, y)(x^0 - x)$，从而定理得证. ▯

定义 8.1.2 设集合 $X \subset \mathscr{B}$ 非空，$\psi: X \to 2^{\mathscr{D}}$, $x \mapsto \psi(x)$ 是集

值映射. 点 $x^0 \in X$, $K \subset \mathscr{D}$ 是闭凸锥. 若对任意的 $y^0 \in \psi(x^0)$ 有 $\psi(x) \subset y^0 + K (\forall x \in X)$, 则称 x^0 是 ψ(在 X 上)的最小点.

定理 8.1.8 设 $x^0 \in X$, $y^0 \in \psi(x^0) \in \mathscr{D}$, K^* 是 K 的对偶锥.

(1) x^0 是 ψ 的最小点当且仅当

$$\psi^T(x^0, y^0)(u) \subset K \quad \forall u \in X. \tag{8.1.3}$$

(2) x^0 是 ψ 的最小点当且仅当

$$\theta \in \partial^T \psi(x^0, y^0)(q^*) \quad \forall q^* \in K^*. \tag{8.1.4}$$

证明 (1) 设(8.1.3)成立, 则对任意的 $x \in X$, 由定理 8.1.7 得

$$\psi(x) \subset y^0 + \psi^T(x^0, y^0)(x - x^0) \subset y^0 + K.$$

因此, 由定义 8.1.2 可知, x^0 是 ψ 的最小点.

反之, 设 x^0 是 ψ 的最小点, 则对任意的 $v \in \psi^T(x^0, y^0)(u)$ 和任意的 $\varepsilon > 0$, 存在 $\bar{u} \in u + \varepsilon B$ ($B \subset \mathscr{B}$ 是单位球)和 $t > 0$, 使得

$$v \in \frac{\psi(x^0 + t\bar{u}) - y^0}{t}.$$

由于 K 是闭的, 令 $\varepsilon \to 0$ 得 $v \in K$, 即(8.1.3)成立.

(2) 我们证明(8.1.3)和(8.1.4)等价. 事实上, 对任意的 $u \in \operatorname{dom} \psi^T(x^0, y^0)$, $v \in \psi^T(x^0, y^0) \subset K$, $q^* \in K^*$, 有

$$\langle \theta, u \rangle - \langle q^*, v \rangle \geqslant 0.$$

由定理 8.1.1 即得 $\theta \in \partial^T \psi(x^0, y^0)(q^*)$. 反之, 设 $\theta \in \partial^T \psi(x^0, y^0)(q^*) (\forall q^* \in K^*)$, 对于 $v \in \psi^T(x^0, y^0)(u)$, 则有

$$\langle \theta, u \rangle - \langle q^*, v \rangle \leqslant 0, \quad \forall q^* \in K^*,$$

故 $v \in K$. 于是, 定理得证. ▯

定理 8.1.9 设 $S \subset \mathscr{B}$ 和 $M \subset \mathscr{D}$ 是非空闭凸集, $\mathscr{A} \in$

$L(\mathscr{B}, \mathscr{D})$ 是线性算子. 定义集值映射 $\psi: \mathscr{B} \to 2^{\mathscr{D}}$,

$$\psi(x) = \begin{cases} \mathscr{A}(x) - M, & x \in S; \\ \varnothing, & x \notin S, \end{cases}$$

则对任意的 $x \in S$ 和 $y \in \mathscr{A}(x) - M$, 有

$$\psi^T(x, y)(u) = \begin{cases} \mathscr{A}(x) - T_M(\mathscr{A}(x) - y), & u \in T_S(x); \\ \varnothing, & u \notin T_S(x) \end{cases}$$

和

$$\partial^T \psi(x, y)(q^*) = \begin{cases} \mathscr{A}^*(q^*) - N_S(x), & q \in N_M(\mathscr{A}(x) - y); \\ \varnothing, & q \notin N_M(\mathscr{A}(x) - y). \end{cases}$$

证明　设 $v \in \psi^T(x, y)(u)$, 由定义 8.1.1 知存在序列 $\{u^k\}$, $\{v^k\}$ 和 $\{t_k\}(t_k > 0, k = 1, 2, \cdots)$, $u^k \to u$, $v^k \to v$, $t_k \to 0$, 使得

$$y + t_k v^k \in \psi(x + t_k u^k), \ k = 1, 2, \cdots.$$

也就是说有 $x + t_k u^k \in S(k = 1, 2, \cdots)$, 即 $u \in T_S(x)$, 并且有

$$\mathscr{A}(x) - y + t_k(\mathscr{A}(u^k) - v^k) \in M, \ k = 1, 2, \cdots.$$

因此,我们有 $\mathscr{A}(u) - v \in T_M(\mathscr{A}(x) - y)$.

反之,设 $u \in T_S(x)$ 和 $v \in \mathscr{A}(u) - T_M(\mathscr{A}(x) - y)$, 则存在 $u^k \to u$, $w^k \to \mathscr{A}(u) - v$ 和 $t_k^1, t_k^2 > 0$, 使得

$$x + t_k^1 u^k \in S, \ k = 1, 2, \cdots,$$

$$\mathscr{A}(x) - y + t_k^2 w^k \in M, \ k = 1, 2, \cdots.$$

设 $t^k = \min(t_k^1, t_k^2)$ 和 $v^k = \mathscr{A}(u^k) - w^k$, 则有 $v^k \to v$. 从上式得

$$y + t_k v^k \in \psi(x + t_k u^k), \ k = 1, 2, \cdots,$$

故有 $v \in \psi^T(x, y)(u)$. 由定义 8.1.1 可知, $p^* \in \psi^T(x, y)(q^*)$ 当且仅当

$$\sup_{u \in T_S(x)} \sup_{w \in T_M(\mathscr{A}(x) - y)} \left[\langle p^*, u\rangle + \langle -q^*, \mathscr{A}(u) - w\rangle\right]$$

$$= \sup_{u \in T_S(x)} \langle p^* - \mathscr{A}^*(q^*), u \rangle + \sup_{w \in T_M(\mathscr{A}(x)-y)} \langle q^*, w \rangle \leqslant 0.$$

此即 $p^* - \mathscr{A}^*(q^*) \in N_S(x)$ 和 $q^* \in N_M(\mathscr{A}(x) - y)$，从而定理得证. ▯

下面研究集值映射的切导数和切微分的计算问题.

再设 $\mathscr{B}_1$ 是 Banach 空间，$\psi: \mathscr{B} \to 2^{\mathscr{D}}$，$x \mapsto \psi(x)$ 是集值映射.

定义 8.1.3 若集值映射 ψ 的图象 graph ψ 是闭的或凸的，则分别称 ψ 在 $\mathscr{B}$ 上是闭的或凸的.

定理 8.1.10 设集值映射 ψ 在 $\mathscr{B}$ 上是闭凸的，$\mathscr{A} \in L(\mathscr{B}_1, \mathscr{B})$. 若

$$\boldsymbol{0} \in \operatorname{int}(\operatorname{Im} \mathscr{A} - \operatorname{dom} \psi),$$

则有

$$(\psi\mathscr{A})^T(z, y) = \psi^T\big(\mathscr{A}(z), y\big)\mathscr{A}, \tag{8.1.5}$$

$$\partial^T(\psi\mathscr{A})(z, y) = \mathscr{A}^* \partial^T\psi\big(\mathscr{A}(z), y\big). \tag{8.1.6}$$

证明 作集值映射 $g: \mathscr{B}_1 \to 2^{\mathscr{D}}$，$z \mapsto g(z) = \psi\mathscr{A}(z)$. 由于 $\mathscr{A}$ 是线性算子，ψ 在 $\mathscr{B}$ 上是闭凸的，因此 g 在 $\mathscr{B}_1$ 上是闭凸的. graph $g = (\mathscr{A} \times 1)^{-1}$graph ψ，于是根据定理 5.1.13，我们有

$$T_{\mathrm{graph}g}(z, y) = (\mathscr{A} \times 1)^{-1} T_{\mathrm{graph}g}\big(\mathscr{A}(z), y\big),$$

$$N_{\mathrm{graph}g}(z, y) = (\mathscr{A}^* \times 1) N_{\mathrm{graph}g}\big(\mathscr{A}(z), y\big).$$

从上式即可推得(8.1.5)和(8.1.6). ▯

定理 8.1.11 设集值映射 ψ 在 $\mathscr{B}$ 上是凸的，$\mathscr{A} \in L(\mathscr{D}, \mathscr{B}_1)$，作集值映射 $\mathscr{A}\psi: \mathscr{B} \to 2^{\mathscr{B}_1}$，$x \mapsto \mathscr{A}\psi(x) = \bigcup_{y \in \psi(x)} \mathscr{A}(y)$. 若 $\mathscr{A}\psi$ 在 $\mathscr{B}$ 上是凸的，并且 $\boldsymbol{0} \in \operatorname{int}(\operatorname{Im}\mathscr{A} - \operatorname{dom} \psi)$，则

$$(\mathscr{A}\psi)^T\big(x, \mathscr{A}(y)\big) = \operatorname{cl}\big(\mathscr{A}\psi^T(x, y)\big), \tag{8.1.7}$$

$$\partial^T(\mathscr{A}\psi)\big(x, \mathscr{A}(y)\big) = \partial^T\psi(x, y)\mathscr{A}^*. \tag{8.1.8}$$

证明 由 graph$(\mathscr{A}\psi) = (1 \times \mathscr{A})$graph ψ 成立. 于是，由定

理 5.1.11 得

$$T_{\mathrm{graph}(\mathscr{A}\psi)}\big(x, \mathscr{A}(y)\big) = \mathrm{cl}\big((1 \times \mathscr{A})T_{\mathrm{graph}\psi}(x, y)\big),$$

$$N_{\mathrm{graph}(\mathscr{A}\psi)}\big(x, \mathscr{A}(y)\big) = \big((1 \times \mathscr{A})^{*-1}N_{\mathrm{graph}\psi}(x, y)\big).$$

从上面两式，即可得知(8.1.7)和(8.1.8)成立. 定理得证. ▯

再设 $\mathscr{B}_2$ 是 Banach 空间.

定理 8.1.12　设集值映射 ψ 在 $\mathscr{B}$ 上是闭凸的，$\mathscr{A} \in L(\mathscr{B}_2, \mathscr{B})$，$\mathscr{A}_1 \in L(\mathscr{D}, \mathscr{B}_1)$，集值映射 $\mathscr{A}_1\psi: \mathscr{B} \to 2^{\mathscr{B}_1}$ 是凸的，又 $u \in \mathscr{B}_2$，$y \in \psi(\mathscr{A}(u))$. 若 $\theta \in \mathrm{int}(\mathrm{Im}\mathscr{A} - \mathrm{dom}\,\psi)$，则

$$(\mathscr{A}_1\psi\mathscr{A})^T\big(u, \mathscr{A}_1(y)\big) = \mathrm{cl}\big(\mathscr{A}_1\psi^T(\mathscr{A}(u), y)\mathscr{A}\big),$$

$$\partial^T(\mathscr{A}_1\psi\mathscr{A})\big(u, \mathscr{A}_1(y)\big) = \mathscr{A}^*\partial^T\psi\big(\mathscr{A}(u), y\big)\mathscr{A}_1^*.$$

证明　由定理 8.1.10 和定理 8.1.11 可得证. ▯

定理 8.1.13　设 $\psi_1: \mathscr{B} \to 2^{\mathscr{B}_1}$ 和 $\psi_2: \mathscr{D} \to 2^{\mathscr{B}_1}$ 是集值映射，$\mathscr{A} \in L(\mathscr{B}, \mathscr{D})$，graph$\psi_1$ 和 graph ψ_2 是闭凸的，并且 $\theta \in \mathrm{int}(\mathscr{A}\mathrm{dom}\,\psi_1 - \mathrm{dom}\psi_2)$. 若 $x \in \mathrm{dom}\,\psi_1 \cap \mathscr{A}^{-1}\mathrm{dom}\,\psi_2$，$y \in \psi_1(x)$，$z \in \psi_2(\mathscr{A}(x))$，则

$$(\psi_1 + \psi_2\mathscr{A})^T(x, y+z) = \mathrm{cl}\big(\psi_1^T(x, y) + \psi_2^T(\mathscr{A}(x), z)\mathscr{A}\big),$$

$$\partial^T(\psi_1 + \psi_2\mathscr{A})(x, y+z) = \partial^T\psi_1(x, y) + \mathscr{A}^*\partial^T\psi_2(\mathscr{A}(x), z).$$

证明　集值映射 $\psi_1 + \psi_2\mathscr{A}: \mathscr{B} \to 2^{\mathscr{B}_1}$ 可以分解成以下三个映射的复合：

$$\mathscr{A}_1: x \in \mathscr{B} \to \big(x, \mathscr{A}(x)\big) \in \mathscr{B} \times \mathscr{D},$$

$$\psi_3: (x, y) \in \mathscr{B} \times \mathscr{D} \to \psi_1(x) \times \psi_2(y) \subset \mathscr{B}_1 \times \mathscr{B}_1,$$

$$A_2: (z_1, z_2) \in \mathscr{B}_1 \times \mathscr{B}_1 \to z_1 + z_2 \in \mathscr{B}_1.$$

也即 $\psi_1 + \psi_2\mathscr{A} = \mathscr{A}_2\psi_3\mathscr{A}_1$，并且有

$$\boldsymbol{\theta} \in \operatorname{int}(\operatorname{Im} \mathscr{A}_1 - \operatorname{dom} \psi_2).$$

设 $\psi_3^T(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}_1, \boldsymbol{z}_2) = \psi_1^T(\boldsymbol{x}, \boldsymbol{z}_1) \times \psi_2^T(\boldsymbol{y}, \boldsymbol{z}_2)$，由定理 8.1.12 即得结论. □

推论 8.1.14 设 graph ψ_1 和 graph ψ_2 是凸的，$\boldsymbol{\theta} \in \operatorname{int}(\operatorname{dom} \psi_1 - \operatorname{dom} \psi_2)$，$\boldsymbol{x} \in \operatorname{dom} \psi_1 \cap \operatorname{dom} \psi_2$，$\boldsymbol{y} \in \psi_1(\boldsymbol{x})$，$\boldsymbol{z} \in \psi_2(\boldsymbol{x})$，则

$$(\psi_1 + \psi_2)^T(\boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z}) = \operatorname{cl}\{\psi_1^T(\boldsymbol{x}, \boldsymbol{y}) + \psi_2(\boldsymbol{x}, \boldsymbol{z})\},$$

$$\partial^T(\psi_1 + \psi_2)(\boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z}) = \partial^T\psi_1(\boldsymbol{x}, \boldsymbol{y}) + \partial^T\psi_2(\boldsymbol{x}, \boldsymbol{z}).$$

证明 由定理 8.1.13 可得证. □

推论 8.1.15 设集值映射 ψ 在 $\mathscr{B}$ 上是闭凸的，$S \subset \mathscr{B}$ 是闭凸集，集值映射 $\theta_S : S \to 2^{\mathscr{D}}$，

$$\theta_S(\boldsymbol{x}) = \begin{cases} 0, & \boldsymbol{x} \in S; \\ \varnothing, & \boldsymbol{x} \notin S. \end{cases}$$

记 $\psi|_S = \psi + \theta_S$，若 $\boldsymbol{\theta} \in \operatorname{int}(\operatorname{dom} \psi - S)$，则对任意的 $\boldsymbol{x} \in S$ 和 $\boldsymbol{y} \in \psi(\boldsymbol{x})$，有

$$(\psi|_S)^T(\boldsymbol{x}, \boldsymbol{y}) = \psi^T(\boldsymbol{x}, \boldsymbol{y}) + \theta_S(\boldsymbol{x}),$$

$$\partial^T(\psi|_S)(\boldsymbol{x}, \boldsymbol{y})(\boldsymbol{q}) = \partial^T\psi(\boldsymbol{x}, \boldsymbol{y})(\boldsymbol{q}) + N_S(\boldsymbol{x}).$$

证明 在推论 8.1.14 中令 $\psi_1 = \psi$，$\psi_2 = \theta_S$，即可得证. □

推论 8.1.16 设集值映射 ψ 在 $\mathscr{B}$ 上是闭凸的，$K \subset \mathscr{D}$ 是闭凸锥，记集合 $\psi_+(\boldsymbol{x}) = \psi(\boldsymbol{x}) + K(\forall \boldsymbol{x} \in \mathscr{B})$，则对任意 $\boldsymbol{x} \in \operatorname{dom} \psi$ 和任意 $\boldsymbol{y} \in \psi(\boldsymbol{x})$，有

$$\psi_+^T(\boldsymbol{x}, \boldsymbol{y})(\boldsymbol{u}) = \psi^T(\boldsymbol{x}, \boldsymbol{y})(\boldsymbol{u}) + K,$$

$$\partial^T\psi_+(\boldsymbol{x}, \boldsymbol{y})(\boldsymbol{q}^*) = \begin{cases} \partial^T\psi(\boldsymbol{x}, \boldsymbol{y})(\boldsymbol{q}^*), & \boldsymbol{q}^* \in K^*; \\ \varnothing, & \boldsymbol{q}^* \notin K^*. \end{cases}$$

证明 在推论 8.1.14 中令 $\psi_1 = \psi$，$\psi_2 = K$，即可得证. □

定理 8.1.17　设集值映射 $\psi_i:\mathscr{B}\to 2^{\mathscr{D}}(i=1,\cdots,m)$ 是闭凸的. 若有 $x^0\in\bigcap_{i=1}^{m}\mathrm{int}(\mathrm{dom}\psi_i)$，并且存在 $\gamma>0$，使得

$$\bigcap_{i=1}^{m}\big(\psi_i(x^0+u_i)-v_i\big)\neq\varnothing\quad\forall(u_i,v_i)\in\gamma(B\times B),$$

则对任意的 $x\in\bigcap_{i=1}^{m}\mathrm{dom}\,\psi_i$ 和任意的 $y\in\bigcap_{i=1}^{m}\psi_i(x)$，有

$$\Big(\bigcap_{i=1}^{m}\psi_i\Big)^T(x,y)=\bigcap_{i=1}^{m}\psi_i^T(x,y).$$

证明　由于 $\mathrm{graph}\Big(\bigcap_{i=1}^{m}\psi_i\Big)=\bigcap_{i=1}^{m}\mathrm{graph}\,\psi_i$，由已设可知存在 $\delta>0$，使得

$$\bigcap_{i=1}^{m}\big(\mathrm{graph}\,\psi_i-(u_i,v_i)\big)\neq\varnothing\quad\forall(u_i,v_i)\in\delta(B\times B).$$

事实上，取 $x^0\in\bigcap_{i=1}^{m}\mathrm{dom}\,\psi_i$ 和 $\delta\leqslant\gamma$，使得

$$x^0+\delta B\subset\bigcap_{i=1}^{m}\mathrm{dom}\,\psi_i,$$

选择 $y^0\in\bigcap_{i=1}^{m}\big(\psi_i(x^0+u_i)-v_i\big)$，则由已知得

$$(x^0,y^0)\in\bigcap_{i=1}^{m}\big(\mathrm{graph}\,\psi_i-(u_i,v_i)\big).$$

于是，由定理 5.1.17 即得定理的结论. □

§ 8.2　D - 切导数

现在，利用 D-切锥（相依切锥）来定义一般集值映射的 D-切导数.

设 $\mathscr{B}$ 和 $\mathscr{D}$ 是 Banach 空间，$\psi:\mathscr{B}\to 2^{\mathscr{D}}$，$x\mapsto\psi(x)$ 是集值映射，点 $x^0\in\mathscr{B}$，$y^0\in\psi(x^0)$，$T^D_{\mathrm{graph}\psi}(x^0,y^0)$ 是 ψ 的图象 graph ψ 在点 (x^0,y^0) 处的 D-切锥.

定义 8.2.1 设点 $\boldsymbol{x}^0 \in \mathscr{B}$, $\boldsymbol{y}^0 \in \psi(\boldsymbol{x}^0)$, $\boldsymbol{u} \in \mathscr{B}$, $\psi^D(\boldsymbol{x}, \boldsymbol{y}): \mathscr{B} \to 2^{\mathscr{D}}$, $u \mapsto \psi^D(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{u})$ 是集值映射. 若 $\boldsymbol{v} \in \psi^D(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{u})$ 当且仅当 $(\boldsymbol{u}, \boldsymbol{v}) \in T^D_{\mathrm{graph}\psi}(\boldsymbol{x}^0, \boldsymbol{y}^0)$, 则称 $\psi^D(\boldsymbol{x}^0, \boldsymbol{y}^0)$是集值映射 ψ 在点 $(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的 D-切导数.

注 8.2.1 由注 5.2.1 我们知道, $\boldsymbol{v}^0 \in \psi^D(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{u}^0)$ 当且仅当存在 $t_k \to 0^+$, $\boldsymbol{u}^k \to \boldsymbol{u}^0$ 和 $\boldsymbol{v}^k \to \boldsymbol{v}^0$, 使得 $\boldsymbol{v}^k \in (\psi(\boldsymbol{x}^0 + t_k\boldsymbol{u}^k) - \boldsymbol{y}^0)/t_k$ $(k = 1, 2, \cdots.)$, 或者

$$\liminf_{\substack{\boldsymbol{u} \to \boldsymbol{u}^0 \\ t \to 0^-}} d\left(\boldsymbol{v}^0, \frac{\psi(\boldsymbol{x}^0 + t\boldsymbol{u}) - \boldsymbol{y}^0}{t}\right) = 0. \qquad (8.2.1)$$

特别是,当 ψ 是单值映射时,上式可写作

$$\liminf_{\substack{\boldsymbol{u} \to \boldsymbol{u}^0 \\ t \to 0^+}} \frac{\|\psi(\boldsymbol{x}^0 + t\boldsymbol{u}) - \psi(\boldsymbol{x}^0) - t\,\boldsymbol{v}^0\|}{t} = 0.$$

进一步,若单值映射 ψ 在点 $\boldsymbol{x}^0$ 处是严格可微的,则

$$\psi^D(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{u}^0) = D_S\psi(\boldsymbol{x}^0)\boldsymbol{u}^0.$$

当 graph ψ 是凸的,则有 $\psi^D(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{u}^0) = \psi^D(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{u}^0)$.

注 8.2.2 集值映射 ψ 在点 $\boldsymbol{x}^0$ 处是 Lipschitz 的,即存在常数 $\gamma > 0$ 使得

$$\psi(\boldsymbol{x}^0 + t\boldsymbol{u}) - \boldsymbol{y}^0 \subset \psi(\boldsymbol{x}^0 + t\boldsymbol{u}^0) - \boldsymbol{y}^0 + \gamma t\|\boldsymbol{u} - \boldsymbol{u}^0\| B.$$

定理 8.2.1 设 $\boldsymbol{x} \in \mathrm{int}(\mathrm{dom}\,\psi)$. 若集值映射 ψ 在 $\boldsymbol{x}$ 附近是 Lipschitz 的,则 $\boldsymbol{v} \in \psi^D(\boldsymbol{x}, \boldsymbol{y})(\boldsymbol{u})$ 当且仅当

$$\liminf_{t \to 0^+} d\left(\boldsymbol{v}, \frac{\psi(\boldsymbol{x} + t\boldsymbol{u}) - \boldsymbol{y}}{t}\right) = 0. \qquad (8.2.2)$$

当 $\mathscr{D}$ 是有限维空间时,则有 $\mathrm{dom}(\psi^D(\boldsymbol{x}, \boldsymbol{y})) = \mathscr{B}$.

证明 必要性. 由于 ψ 在 $\boldsymbol{x}$ 处是 Lipschitz 的,则有

$$\psi(\boldsymbol{x} + t\boldsymbol{u}') - \boldsymbol{y} \subset \psi(\boldsymbol{x} + t\boldsymbol{u}) - \boldsymbol{y} + \gamma t\|\boldsymbol{u}' - \boldsymbol{u}\| B.$$

令 $t \to 0^+$, $u^t \to u$ 由(8.1.1)即得(8.2.2).

充分性. 设取任意的 $u \in \mathscr{B}$, 则由(8.2.2)对充分小的 t, 有

$$y \in \psi(x) \subset \psi(x+tu) + \gamma t \|u\| B.$$

故存在 $v^t \in \psi(x+tu)$ 使得 $\frac{(v^t - y)}{t} \in \gamma \|u\| B$. 由于 $\gamma\|u\|B$ 是紧的, 故存在收敛子列 $\{(v^t_k - y)/t_k\}$, $(v^t_k - y)/t_k \to v$ $(k \to \infty)$, 由定义 8.2.1 便得 $v \in \psi^D(x, y)(u)$. 当 ψ 是有限维空间时, 由上述结果易知第 2 个结论成立. ▯

定理 8.2.2 设 $X \subset \mathscr{B}$ 是非空开集. 若映射 $\varphi: X \to \mathscr{D}$ 是严格可微的, 集合 $S \subset X$ 非空, $x \in S$, 则

$$\varphi|_S^D(x)(u) = \begin{cases} D_S\varphi(x)u, & u \in T^D(x); \\ \varnothing, & u \notin T^D(x). \end{cases}$$

证明 当 $u \in T_S^D(x)$ 时, 则存在 $t_k \to 0^+$, $u^k \to u(k \to \infty)$, 使得 $x + t_k u^k \in S$. 因为

$$\varphi|_S(x + t_k u^k) = \varphi(x + t_k u^k)$$

$$= \varphi(x) + t_k\left(D_S\varphi(x)u^k + o(t_k)\right),$$

令 $v^k = D_S\varphi(x)u^k + o(t_k)$, 则 $v^k \to D_S\varphi(x_0)u_0$, 并且有

$$v^k \in \{\varphi|_S(x + t_k u^k) - \varphi|_S(x)\}/t_k.$$

由此, 得到 $\varphi^D\left(x, \varphi(x)\right)(u) = D_S\varphi(x)u$. ▯

设 $\mathscr{B}_1$ 是 Banach 空间.

定理 8.2.3 设 $x \in \mathscr{B}$, $y \in \psi(x)$, Ω 是 $\operatorname{Im}\psi$ 的一个开邻域. 若映射 $\varphi: \Omega \to \mathscr{B}_1$ 是严格可微的, 则

$$D_S\varphi(y)\psi^D(x, y)(u) \subset (\varphi\psi)^D\left(x, \varphi(y)\right)(u)$$

$$\forall u \in \mathscr{B}. \tag{8.2.3}$$

再若 ψ 在 $x \in \operatorname{int}(\operatorname{dom}\psi)$ 处是 Lipschitz 的, 并且是紧值的, $\mathscr{D}$ 是有限维空间, 则(8.2.3)成为等式.

证明 作映射 $(1\times\varphi):(x,y)\in\mathscr{B}\times\Omega\to(x,\varphi(y))\in\mathscr{D}\times\mathscr{B}_1$ 对集值映射 $\theta=\varphi\psi:\mathscr{B}\to2^{\mathscr{B}_1}$, $x\mapsto\varphi\psi(x)\subset\mathscr{B}_1$ 有

$$\operatorname{graph}\theta=(1\times\varphi)\operatorname{graph}\psi.$$

据此,由定理 4.2.6 即可得证.

为证明(8.2.3)成为等式,设 $w\in(\varphi\psi)^D(x,\varphi(y))(u)$. 则存在 $t_k\to0$, $u^k\to u$ 和 $w^k\to w(k\to\infty)$, 使得

$$t_kw^k\in\varphi(\psi(x+t_ku^k))-\varphi(y),\ k=1,2,\cdots.$$

因此,存在 $v^k(k=1,2,\cdots)\in(\psi(x+t_ku^k)-y)/t_k$, 使得

$$t_kw^k=\varphi(y+t_kv^k)-\varphi(y),\ k=1,2,\cdots.$$

因为 ψ 在 x 处是Lipschitz的,故有 $v^k\in\psi(x)-y+t^k\|u^k\|B$. 再由 $\mathscr{D}$ 是有限维空间和 ψ 是紧值的,所以 $\{v^k\}$ 中存在收敛子列,不妨设为 $v^k\to v$. 于是有 $v\in\psi^D(x,y)(u)$, 并且 $w^k\to D_S\varphi(y)v=w$. 由此得到 $(\varphi\psi)^D(x,\varphi(y))(u)\subset D_S\varphi(y)\psi^D(x,y)(u)$. □

定理 8.2.4 设 $\psi_i:\mathscr{B}\to2^{\mathscr{D}}$, $x\mapsto\psi_i(x)$ $(i=1,2,\cdots,m)$ 是集值映射, $x\in\operatorname{dom}\psi_i(i=1,2,\cdots,m)$. 若 $y\in\bigcap_{i=1}^{m}\operatorname{Im}\psi_i(x)$, 则

$$\left(\bigcup_{i=1}^{m}\psi_i\right)^D(x,y)(u)=\bigcup_{i=1}^{m}\psi_i^D(x,y)(u).$$

证明 根据定理 4.2.5 和定理 4.2.7,可以得证. □

§8.3 C-切导数和C-切微分

本节将用C-切锥,引进一般集值映射的C-切导数和C-切微分.

设 $\mathscr{B}$ 和 $\mathscr{D}$ 是 Banach 空间, $\mathscr{B}^*$ 和 $\mathscr{D}^*$ 分别是 $\mathscr{B}$ 和 $\mathscr{D}$ 的对偶空间, $\psi:\mathscr{B}\to2^{\mathscr{D}}$, $x\mapsto\psi(x)$ 是集值映射, $(x^0,y^0)\in\operatorname{graph}\psi$, $T^C_{\operatorname{graph}\psi}(x^0,y^0)$ 是 graph ψ 在点 (x^0,y^0) 处的C-切锥.

定义 8.3.1 设点 $x^0\in\mathscr{B}$, $y^0\in\psi(x^0)$, $u\in\mathscr{B}$, 考虑集值映

射 $\psi^C(x^0, y^0): \mathscr{B} \to 2^{\mathscr{D}}$, $u \mapsto \psi^C(x^0, y^0)(u)$. 若 $v \in \psi^C(x^0, y^0)(u)$ 当且仅当

$$(u, v) \in T_{\text{graph}\psi}(x^0, y^0),$$

则称 $\psi^C(x^0, y^0)$是集值映射 ψ 在点(x^0, y^0)处的 C-切导数. 考虑集值映射 $\partial^C\psi(x^0, y^0): \mathscr{D}^* \to 2^{\mathscr{B}^*}$, $q_0^* \mapsto \partial^C\psi(x^0, y^0)(q_0^*)$. 若 $p_0^* \in \partial^C\psi(x^0, y^0)(q_0^*)$ 当且仅当

$$\langle p_0^*, u\rangle \leqslant \langle q_0^*, v\rangle \quad \forall u \in \mathscr{B}, \forall v \in \psi^C(x^0, y^0)(u),$$

则称 $\partial^C\psi(x^0, y^0)$是集值映射 ψ 在点(x^0, y^0)处的 C-切微分.

由定义 5.2.2 可以推得以下性质.

定理 8.3.1 设 $x^0 \in \mathscr{B}$, $y^0 \in \psi(x^0)$, 则 $v^0 \in \psi^C(x^0, y^0)(u^0)$ 与下面两命题等价:

(1) 若对任意的 ε_1, $\varepsilon_2 > 0$, 存在 α, $\beta > 0$, 使得 $(x, y) \in B_{\text{graph}\psi}(x^0, y^0, \alpha)$, 则对任意的 $t(\alpha, \beta)$, 存在 $u \in u^0 + \varepsilon_1 B$, $v \in v^0 + \varepsilon_2 B$, 使得

$$v \in \big(\psi(x + tu) - y\big)/t.$$

(2) 对所有的收敛于$(x^0, y^0, 0)$的序列 $(x^k, y^k, t_k) \in \text{graph}\,\psi \times (0, +\infty)$, 存在序列$\{u^k\}$和$\{v^k\}$, $u^k \to u^0$, $v^k \to v^0$ $(k \to \infty)$, 使得

$$y^k + t_k v^k \in \psi(x^k + t_k u^k), \quad k = 1, 2, \cdots.$$

定理 8.3.2 设 $x^0 \in \mathscr{B}$, $y^0 \in \psi(x^0)$. 若 ψ 在 $x^0 \in \text{int}(\text{dom}\,\psi)$ 处是 Lipschitz 的, 则 $v^0 \in \psi^C(x^0, y^0)(u^0)$ 当且仅当

$$\lim_{\substack{x \to x^0 \\ t \to 0^+}} d\left(v^0, \frac{\psi(x + tu^0) - y^0}{t}\right) = 0.$$

定理 8.3.3 设 $x^0 \in R^n$, f 是单值连续可微函数, $y^0 = f(x^0)$, 则

$$f^C\big(x^0, f(x^0)\big) = \nabla f(x^0).$$

由定义 7.3.1,定义 7.2.1 以及定理 7.2.3,可以得到 C-弱方向导数和 C-弱次微分与 C-切导数和 C-切微分之间的关系.

定理 8.3.4　设 $\boldsymbol{x}^0 \in \mathscr{B}$, $\boldsymbol{y}^0 \in \psi(\boldsymbol{x}^0)$, $K \subset \mathscr{D}$ 是内部非空的尖闭凸锥, $\boldsymbol{p} \in \mathrm{int}K$, $\boldsymbol{d} \in \mathscr{B}$, $\Psi(\boldsymbol{x}^0, \boldsymbol{y}^0, \boldsymbol{d})$ 由(7.2.1)确定, $\hat{D} = \mathscr{B}$. 又设 $\Psi(\boldsymbol{x}^0, \boldsymbol{y}^0, \boldsymbol{d})$ 是 K-有界的.

(1) $\psi^C(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{d}) \subset \psi_w^C(\boldsymbol{x}^0, \boldsymbol{y}^0; \boldsymbol{d})$.

(2) 若 $\boldsymbol{x}^* \in \partial^C\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{q}^*)$, $\boldsymbol{q}^* \in K^* \setminus \{\boldsymbol{0}\}$, 则

$$\boldsymbol{x}^*\{\langle \boldsymbol{q}^*, \boldsymbol{p}\rangle\}^{-1} \in \partial_w^C\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)_p.$$

(3) 若 $\boldsymbol{x}^* \in \partial_w^C\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)_p$, 则存在 $\boldsymbol{q}^* \in K^* \setminus \{\boldsymbol{0}\}$ 使得

$$\langle \boldsymbol{q}^*, \boldsymbol{p}\rangle \boldsymbol{x}^* \in \partial^C\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{q}^*).$$

定理 8.3.5　设 $\Omega \subset \mathscr{B}$ 是非空开集. 若函数 $f: \Omega \to R$ 在 $\boldsymbol{x} \in \Omega$ 处是连续可微的,集合 $S \subset \mathscr{B}$, $\boldsymbol{x} \in S$, 则

$$f|_S^C(\boldsymbol{x}, f(\boldsymbol{x}))(\boldsymbol{u}) = \begin{cases} \nabla f(\boldsymbol{x})\boldsymbol{u}, & \boldsymbol{u} \in T_S(\boldsymbol{x}); \\ \varnothing, & \boldsymbol{u} \notin T_S(\boldsymbol{x}). \end{cases}$$

证明　设序列 $\{\boldsymbol{x}^k, t_k\} \in S \times (0, 1)$ 收敛于$(\boldsymbol{x}, 0)$,当 $\boldsymbol{u} \in T_S(\boldsymbol{x})$ 时,则存在序列 $\{\boldsymbol{u}^k\}$, $\boldsymbol{u}^k \to \boldsymbol{u}(k \to \infty)$, 使得

$$\boldsymbol{x}^k + t_k\boldsymbol{u}^k \in S, \ k = 1, 2, \cdots,$$

有

$$\begin{aligned} f|_S(\boldsymbol{x}^k + t_k\boldsymbol{u}^k) &= f(\boldsymbol{x}^k + t_k\boldsymbol{u}^k) \\ &= f(\boldsymbol{x}^k) + t_k\left(\nabla f(\boldsymbol{x}^k)\boldsymbol{u}^k + O(t_k)\right). \end{aligned}$$

因为 f 是连续可微的,所以序列 $\{\boldsymbol{v}^k\} = \left\{\nabla f(\boldsymbol{x}^k)\boldsymbol{u}^k + O(t_k)\right\}$ 收敛于 $\nabla f(\boldsymbol{x})\boldsymbol{u}$,由此有

$$f|_S(\boldsymbol{x}^k) + t_k\boldsymbol{v}^k = f|_S(\boldsymbol{x} + t_k\boldsymbol{u}^k), \ k = 1, 2, \cdots.$$

根据定理 8.3.1 即知结论成立. ▯

定理 8.3.6　设 $\Omega \subset \mathscr{B}$ 是开集, $\mathscr{A}: \Omega \to \mathscr{D}$ 是连续可微算子,

$S \subset \Omega$ 和 $Y \subset \mathscr{D}$ 分别是 $\mathscr{B}$ 和 $\mathscr{D}$ 中的闭集. 若集值映射 $\psi: \mathscr{B} \to 2^{\mathscr{D}}$,

$$\psi(x) = \begin{cases} \mathscr{A}(x) - Y, & x \in S; \\ \varnothing, & x \notin S, \end{cases}$$

并且 $(x, y) \in \operatorname{graph} \psi$, 则 $v \in \psi^C(x, y)(u)$ 当且仅当 $u \in T_S(x)$ 和 $v \in \mathscr{A}'_x(u) - T_Y(\mathscr{A}(x) - y)$.

证明　必要性. 设 $v \in \psi^C(x, y)(u)$, 取收敛于 $\{x, \mathscr{A}(x - y), 0\}$ 的序列 $\{x^k, z^k, t_k\} \in S \times Y \times (0, +\infty)$, 则序列 $y^k = \mathscr{A}(x^k) - z^k$ 收敛于 y. 由定义 8.3.1 可知, 存在序列 $\{u^k\}$ 和 $\{v^k\}$, $u^k \to u$, $v^k \to v$ $(k \to \infty)$, 使得

$$x^k + t_k u^k \in S, \mathscr{A}(x^k + t_k u^k) \in Y + y^k + t_k v^k, k = 1, 2, \cdots.$$

这意味着有 $u \in T_S(x)$ 和

$$w^k = \mathscr{A}(x^k + t_k u^k) - \mathscr{A}(x^k) - v^k \to \mathscr{A}'_x(u) - v,$$

$$z^k + t_k w^k \in Y, k = 1, 2, \cdots.$$

由此, 得 $v \in \mathscr{A}'_x(u) - T_Y(\mathscr{A}(x) - y)$.

充分性. 取收敛于 $(x, y, 0)$ 的序列 $\{x^k, y^k, t_k\} \in \operatorname{graph} \psi \times (0, +\infty)$. 由 $u \in T_S(x)$ 和 $v \in \mathscr{A}'_x(u) - T_S(\mathscr{A}(x) - y)$ 得知存在序列 $\{u^k\}$, $u^k \to u$ $(k \to \infty)$, 使得

$$x^k + t_k u^k \in S, k = 1, 2, \cdots.$$

于是, 存在序列 $\{w^k\}$, $w^k \to \mathscr{A}'_x(u) - v$ $(k \to \infty)$, 使得

$$\mathscr{A}(x^k) - y^k + t_k w^k \in Y, k = 1, 2, \cdots.$$

据此, 则有 $v^k = \{[\mathscr{A}(x^k + t_k u^k) - \mathscr{A}(x^k)]/t^k\} - w^k \to v^0$ $(k \to \infty)$, 使得 $y^k + t_k v^k \in \psi(x^k + t_k u^k)$ $(k = 1, 2, \cdots)$. ▯

定理 8.3.7　设 $\mathscr{B}$ 是 Hilbert 空间, $S \subset \mathscr{B}$ 是闭凸集, $N_S(x)$ 是 S 在点 x 处的法锥, $p \in N_S(x)$, 作集值映射 $N_S: \mathscr{B} \to 2^{\mathscr{B}^*}$, $x \mapsto N_S(x)$. 若单值映射 $\pi_S: x \to \pi_S(x) \in S$, $\pi_S(x)$ 使得下式成立:

$$\pi_S^{-1}(x) = x + N_S(x),$$

则 $q \in N_S^C(x, p)(u)$ 当且仅当 $u \in \pi_S^C(x, p)(u + q)$.

证明 显然，$p \in N_S(x)$ 当且仅当 $x = \pi_S(x + p)$.

必要性. 设 $q \in N_S^C(x, p)(u)$，考虑序列 $\{(y^k, t_k)\} \subset \mathscr{B} \times (0, +\infty)$,

$$(y^k, t_k) \to (x + p, 0), \ k \to \infty.$$

作序列 $\{x^k\} = \{\pi_S(y^k)\}$，则 $x^k \to x = \pi_S(x + p)$，$p^k = y^k - x^k \to p (k \to \infty)$，于是存在序列 $\{u^k\}$ 和 $\{q^k\}$，$u^k \to u$，$q^k \to q (k \to \infty)$，使得

$$p^k + t_k q^k \in N_S(x^k + t_k u^k), \quad k = 1, 2, \cdots,$$

即

$$\pi_S(y^k) + t_k u^k = \pi_S\big(y^k + t_k(q^k + u^k)\big), \quad k = 1, 2, \cdots.$$

因此，我们有 $u \in \pi_S^C(x, p)(u + q)$.

充分性. 设 $u \in \pi_S^C(x, p)(u + q)$，序列 $\{x^k, p^k, t_k\} \in \mathrm{graph} N_S \times (0, +\infty)$ 收敛于 $(x, p, 0)$. 因为 $x^k + p^k$ 收敛于 $x + p$，所以存在序列 $\{u^k\}$ 和 $\{w^k\}$，$u^k \to u$，$w^k \to u + q (k \to \infty)$，使得

$$\begin{aligned} x^k + t_k u^k &= \pi_S(x^k + p^k) + t_k u^k \\ &= \pi_S(x^k + p^k + t_k w^k), \quad k = 1, 2, \cdots. \end{aligned}$$

令 $q^k = w^k - u^k$，则有 $q^k \to q$. 从上式得 $p^k + t_k q^k \in N_S(x^k + t_k u^k)$ $(k = 1, 2, \cdots)$，故 $q \in N_S^C(x, p)(u)$. ▯

定理 8.3.8 设 $\mathscr{B}_0$，$\mathscr{B}$ 和 $\mathscr{D}$ 均为有限维 Euclid 空间，$x' \in \mathscr{B}$，$\psi(x') \subset \mathscr{D}$ 是闭的，$\mathscr{A}: \mathscr{B}_0 \to \mathscr{B}$ 是连续可微映射，$x \in \mathscr{A}^{-1}\mathrm{dom}\psi$ 和 $y \in \psi(x)$. 若

$$\mathrm{Im} J\mathscr{A}(x) - \mathrm{dom}\big\{\psi^C(\mathscr{A}(x), y\big\} = \mathscr{B},$$

则

$$(\psi\mathscr{A})^C(x, y)(u) \supset \psi^C\big(\mathscr{A}(x), y\big\}(J\mathscr{A}(x)u) \quad \forall u \in \mathscr{B},$$

$$\partial^C(\psi\mathscr{A})(x,\ y)(q^*)\subset J\mathscr{A}(x)\partial^C\psi\big(\mathscr{A}(x),\ y\big)(q^*)$$

$$\forall q^*\in\mathscr{D}^*.$$

证明 因为 $\mathrm{graph}\psi\mathscr{A}=(\mathscr{A}\times 1)^{-1}\mathrm{graph}\psi$，由推论 5.2.8 有

$$\big(J\mathscr{A}(x)\times 1\big)^{-1}T^C_{\mathrm{graph}\psi}\big(\mathscr{A}(x),\ y\big)\subset T^C_{\mathrm{graph}\psi\mathscr{A}}(x,\ y).$$

从假设知

$$\mathrm{Im}\big(J\mathscr{A}(x)\times 1\big)-T^C_{\mathrm{graph}\psi}\big(\mathscr{A}(x),\ y\big)=\mathscr{B}\times\mathscr{D},$$

$$\big(\psi^C(\mathscr{A}(x),\ y)J\mathscr{A}(x)\big)=J\mathscr{A}(x)\partial^C\psi\big(\mathscr{A}(x),\ y\big).$$

由此，可得定理结论. □

定理 8.3.9 设集合 $S\subset\mathscr{B}$, $x\in S\cap\mathrm{dom}\psi$, $y\in\psi(x)$，则

$$(\psi|_S)^D(x,\ y)(u)\subset\psi^D(x,\ y)\Big|_{T^D_S(x)}(u).\qquad(8.3.1)$$

若 $\mathscr{B}$ 和 $\mathscr{D}$ 是有限维空间，$\mathrm{graph}\psi$ 和 S 都是闭的，并且

$$T_S(x)-\mathrm{dom}\big(\psi^C(x,\ y)\big)=\mathscr{B},\qquad(8.3.2)$$

则

$$\psi^C(x,\ y)\Big|_{T_S(x)}(u)\subset(\psi|_S)^C(x,\ y)(u),$$

$$\partial^C(\psi|_S)(x,\ y)(q^*)\subset\partial^C\psi(x,\ y)(q^*)+N_S(x))\quad\forall q\in\mathscr{D}^*.$$

证明 因为有 $\mathrm{graph}(\psi|_S)=\mathrm{graph}(\psi)(S\times\mathscr{D})$，则

$$T^D_{\mathrm{graph}\psi}|_S(x,\ y)\subset T^D_{\mathrm{graph}\psi}(x,\ y)\cap\big(T^D_S(x)\times\mathscr{D}\big).$$

从上式即可得(8.3.1). 另外，由(8.3.2)得知

$$T^C_{\mathrm{graph}\psi}(x,\ y)-T_S(x)\times\mathscr{D}=\mathscr{B}\times\mathscr{D},$$

从而结论成立. □

定理 8.3.10 设 $\mathscr{B}$ 和 $\mathscr{D}$ 是有限维空间，$\psi_i: \mathscr{B} \to 2^{\mathscr{D}} (i=1, 2, \cdots, m)$ 是集值映射，$(x, y) \in \bigcap_{i=1}^{m} \mathrm{graph}\psi_i$. 若对任何 $(u_i, v_i) \in \mathscr{B} \times \mathscr{D}$ $(i=1, 2, \cdots, m)$，存在$(u, v) \in \mathscr{B} \times \mathscr{D}$，使得

$$v \in \psi_i^C(x, y)(u + u_i) - v_i, \ i = 1, 2, \cdots, m,$$

则

$$\left(\bigcap_{i=1}^{m} \psi_i\right)^C (x, y)(u) \supset \bigcap_{i=1}^{m} \psi_i^C(x, y)(u) \quad \forall u \in \mathscr{B}.$$

证明 由推论 5.2.10 可使定理得证. ▯

§ 8.4 M-次微分

最后，介绍有限维空间中集值映射的另一种广义微分及其性质.

本节考虑集值映射 $\psi: R^n \to 2^{R^m}$，$x \mapsto \psi(x)$，它的图象和核分别为

$$\mathrm{graph}\psi = \{(x, y) \in R^n \times R^m \mid y \in \psi(x)\}$$

和

$$\mathrm{Ker}\psi = \{x \in R^n \mid 0 \in \psi(x)\}.$$

设 $x^0 \in R^n$，记集合

$$\limsup_{x \to x^0} \psi(x) = \{y \in R^m \mid \exists\ x^k \to x^0, \ y^k \to y, \\ k \to \infty; \ y^k \in \psi(x^k), \ k = 1, 2, \cdots. \},$$

并称它是集值映射 ψ 在点 x^0 处的 KP-上极限集.

设集合 $S \subset R^n$ 非空，$x^0 \in S$，记

$$P(x^0, S) = \{w \in \mathrm{cl}S \mid \ \| x^0 - w \| = \inf_{x \in S} \| x - w \|)\}.$$

定义 8.4.1 设集合 $S \subset R^n$ 非空，点 $x^0 \in \mathrm{cl}S$，则称集合

$$N_S^M(x^0) = \lim_{x \to x^0} \sup\left[\text{cone}(x - P(x, S))\right]$$

是集合 S 在点 x^0 处的 M-法锥. 若 $x^0 \notin \text{cl}S$, 则令 $N_S^M(x^0) = \varnothing$.

注 8.4.1　设 $x_1^0 \in S_1$, $x_2^0 \in S_2$, 由定义 8.4.1 我们有

$$N_{S_1 \times S_2}^M((x_1^0, x_2^0)) = N_{S_1}^M(x_1^0) \times N_{S_2}^M(x_2^0).$$

根据定义 5.2.2 的(2),还有关系:

$$N_S^C(x^0) = \text{cl}(\text{co}N_S^M(x^0)) .$$

定义 8.4.2　设点 $(x^0, y^0) \in \text{cl}(\text{graph}\psi)$, 定义集值映射 $\partial\psi^M(x^0, y^0): R^m \to 2^{R^n}$,

$$\partial^M\psi(x^0, y^0)(y^*) = \{x^* \in R^n | (x^*, -y^*) \in N_{\text{graph}\psi}^M(x^0, y^0)\},$$
(8.4.1)

则称 $\partial^M\psi(x^0, y^0)$ 是集值映射 ψ 在点(x^0, y^0)处的 M-次微分. 若 $(x^0, y^0) \notin \text{cl}(\text{graph}\psi)$,令 $\partial^M\psi(x^0, y^0)(y^*) = \varnothing$, 若 ψ 在点 x^0 处是单值的,令 $\partial^M\psi(x^0) = \partial^M\psi(x^0, y^0)$. 若 $\partial^M\psi(x^0, y^0) \neq \varnothing$, 则称 ψ 在点(x^0, y^0)处是 M-次可微的.

由定义 8.4.2 可知, $\partial^M\psi(x^0, y^0)$ 关于 y^* 是正齐次的,并且 $\partial^M\psi(x^0, y^0)(\boldsymbol{0})$ 是闭锥.

例 8.4.1　设集合 $S \subset R^n$ 非空,集合

$$\psi(x) = \begin{cases} 0, & x \in S; \\ \varnothing, & x \notin S. \end{cases}$$

考虑集值映射 $\psi: R^n \to 2^{R^m}$, $x \mapsto \psi(x)$, 由定义 8.4.2 则有

$$\partial^M\psi(x^0)(y^*) = N_S^M(x^0) \quad \forall x^0 \in S, y^* \in R^m.$$

定义 8.4.3　设集合 $S \subset R^n$ 非空,点 $x^0 \in \text{cl}S$, $\epsilon \geq 0$, 则称集合

$$N_S^\epsilon(x^0, \epsilon) = \left\{x^* \in R^n \middle| \limsup_{\substack{x' \to x^0 \\ x' \in S}} \frac{\langle x^*, x' - x^0 \rangle}{\| x' - x^0 \|} \leq \epsilon\right\}$$
(8.4.2)

是集合 S 在点 $\boldsymbol{x}^0$ 处的 F-法锥(Frechet 的 ϵ-法锥). 若 $\boldsymbol{x}^0 \notin \mathrm{cl}S$, 令 $N_S^F(\boldsymbol{x}^0, \epsilon) = \varnothing$, 当 $\epsilon = 0$ 时,记 $N_S^F(\boldsymbol{x}^0) = N_S^F(\boldsymbol{x}^0, 0)$.

注 8.4.2 有下面关系:

$$N_S^F(\boldsymbol{x}^0) = \limsup_{\boldsymbol{x} \to \boldsymbol{x}^0} N_S^F(\boldsymbol{x}) = \limsup_{\boldsymbol{x} \to \boldsymbol{x}^0,\ \epsilon \to 0^+} N_S^F(\boldsymbol{x}, \epsilon). \quad (8.4.3)$$

定义 8.4.4 设点 $\boldsymbol{x}^0 \in R^n$, $\boldsymbol{y}^0 \in \psi(\boldsymbol{x}^0) \subset R^m$.

(1) 称

$$\begin{aligned} &\partial^{MC}\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{y}^*) \\ &= \{\boldsymbol{x}^* \in R^n \mid (\boldsymbol{x}^*, -\boldsymbol{y}^*) \in N_{\mathrm{graph}\psi}^C(\boldsymbol{x}^0, \boldsymbol{y}^0)\} \end{aligned} \quad (8.4.4)$$

是集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的 MC-次微分.

(2) 称

$$\begin{aligned} &\partial^{MF}\psi(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{y}^*) \\ &= \{\boldsymbol{x}^* \in R^n \mid (\boldsymbol{x}^*, -\boldsymbol{y}^*) \in N_{\mathrm{graph}\psi}^F(\boldsymbol{x}^0, \boldsymbol{y}^0)\} \end{aligned} \quad (8.4.5)$$

是集值映射 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处的 MF-次微分.

注 8.4.3 显然,这里由 C-切锥和 F-法锥引进的 MC-次微分和 MF-次微分与定义 8.4.2 中给出的 M-次微分是不一样的.

注 8.4.4 我们知道,当 S 是凸集时,有 $N_S^C(\boldsymbol{x}^0) = N_S^M(\boldsymbol{x}^0)$. 若 $N_S^C(\boldsymbol{x}^0) = N_S^M(\boldsymbol{x}^0)$ 或 $N_S^M(\boldsymbol{x}^0) = N_S^F(\boldsymbol{x}^0)$ 时,则称集合 S 在点 $\boldsymbol{x}^0$ 处是正则的. 显然,若 $\mathrm{graph}\psi$ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处是正则的,则(8.4.1)与(8.4.4)或(8.4.1)与(8.4.5)是一致的,这时称 ψ 在点$(\boldsymbol{x}^0, \boldsymbol{y}^0)$处是可微正则的.

定理 8.4.1 设 $\boldsymbol{x} \in R^n$, $\boldsymbol{y} \in \psi(\boldsymbol{x})$. 若 ψ 在点 $\boldsymbol{x}$ 附近是单值的,在点 $\boldsymbol{x}$ 处是可微的, $J\psi(\boldsymbol{x}) \in R^{m\times n}$, 即

$$\lim_{\boldsymbol{x}', \boldsymbol{x}'' \to \boldsymbol{x}} \frac{(\psi(\boldsymbol{x}') - \psi(\boldsymbol{x}'') - J\psi(\boldsymbol{x})(\boldsymbol{x}' - \boldsymbol{x}''))}{\|\boldsymbol{x}' - \boldsymbol{x}''\|} = \underset{n}{0}.$$

则 ψ 在$(\boldsymbol{x}, \boldsymbol{y})$处是 M-次可微的和正则的,并且有

$$\partial^M\psi(\boldsymbol{x})(\boldsymbol{y}^*) = \{(J\psi(\boldsymbol{x}))^*\boldsymbol{y}^*\} \quad \forall \boldsymbol{y}^* \in R^m.$$

证明 由定理 8.3.3 和定义 8.4.2 不难推得. ▯

定理 8.4.2 设 graphψ 是凸的,则 ψ 在 $(\boldsymbol{x}, \boldsymbol{y}) \in \mathrm{cl}(\mathrm{graph}\psi)$ 处是可微和正则的,并且对任意的 $\boldsymbol{y}^* \in R^m$ 有

$$\partial^M\psi(\boldsymbol{x}, \boldsymbol{y})(\boldsymbol{y}^*)$$
$$= \Big\{\boldsymbol{x}^* \in R^n \,|\, \langle \boldsymbol{x}^*, \boldsymbol{x}\rangle - \langle \boldsymbol{y}^*, \boldsymbol{y}\rangle$$
$$= \sup\big[\langle \boldsymbol{x}^*, \boldsymbol{x}'\rangle - \langle \boldsymbol{y}^*, \boldsymbol{y}'\rangle \quad \forall (\boldsymbol{x}', \boldsymbol{y}') \in \mathrm{graph}\psi\big]\Big\}.$$

证明 由定理 8.3.4 和定理 6.2.11(当 $p = 1$) 可得证. ▯

定义 8.4.5 设点 $\boldsymbol{x}^0 \in R^n$, $\boldsymbol{y}^0 \in \psi(\boldsymbol{x}^0)$, $B \subset R^m$ 是开单位球. 若存在点 $\boldsymbol{x}^0$ 的邻域 $U(\boldsymbol{x}^0)$ 和点 $\boldsymbol{y}^0$ 的邻域 $U(\boldsymbol{y}^0)$ 以及常数 $\gamma \geqslant 0$, 使得

$$\psi(\boldsymbol{x}') \cap U(\boldsymbol{y}^0) \subset \psi(\boldsymbol{x}) + \gamma\,\|\boldsymbol{x}^* - \boldsymbol{x}\|\,B \quad \forall \boldsymbol{x}, \boldsymbol{x}' \in U(\boldsymbol{x}^0),$$

则称集值映射 ψ 在点 $(\boldsymbol{x}^0, \boldsymbol{y}^0) \in \mathrm{graph}\psi$ 处是伪 Lipschitz 的.

定理 8.4.3 设 graphψ 是闭的, $(\boldsymbol{x}, \boldsymbol{y}) \in \mathrm{graph}\psi$, 则 ψ 在点 $(\boldsymbol{x}, \boldsymbol{y})$ 处是伪 Lipschitz 的当且仅当

$$\partial^M\psi(\boldsymbol{x}, \boldsymbol{y})(\boldsymbol{0}) = \{\boldsymbol{0}\}.$$

证明 由定义 8.4.1 和定义 8.4.5 可得证. ▯

设 ψ^{-1} 是 ψ 的逆映射: $\psi^{-1}(\boldsymbol{y}) = \{\boldsymbol{x} \in R^n \,|\, (\boldsymbol{x}, \boldsymbol{y}) \in \mathrm{graph}\psi\}$, 则有

$$\boldsymbol{y}^* \in \partial^M\psi^{-1}(\boldsymbol{y}, \boldsymbol{x})(\boldsymbol{x}^*) \Longleftrightarrow -\boldsymbol{x}^* \in \partial^M\psi(\boldsymbol{x}, \boldsymbol{y})(-\boldsymbol{y}^*).$$

定义 8.4.6 设 $f: R^n \to R \cup \{\pm\infty\}$ 是广义实值函数,点 $\boldsymbol{x}^0 \in \mathrm{dom}f$.

(1) 称集合

$$\partial^M f(\boldsymbol{x}^0) = \partial^M \mathrm{epi}f(\boldsymbol{x}^0, f(\boldsymbol{x}^0))(1)$$
$$= \{\boldsymbol{x}^* \in R^n \,|\, (\boldsymbol{x}^*, -1) \in N^M_{\mathrm{epi}f}(\boldsymbol{x}^0, f(\boldsymbol{x}^0))\}$$

是 f 在点 $\boldsymbol{x}^0$ 处的 M-次微分. 当 $\boldsymbol{x}^0 \notin \mathrm{dom}f$,则令 $\partial^M f(\boldsymbol{x}^0) = \varnothing$.

(2) 称集合

$$\partial^{\infty} f(\boldsymbol{x}^0) = \partial^M \mathrm{epi} f\big(\boldsymbol{x}^0,\ f(\boldsymbol{x}^0)\big)(\boldsymbol{0})$$

$$= \left\{\boldsymbol{x}^* \in R^n \mid (\boldsymbol{x}^*,\ 0) \in N^M_{\mathrm{epi} f}(\boldsymbol{x}^0,\ f(\boldsymbol{x}^0))\right\}$$

是 f 在点 $\boldsymbol{x}^0$ 处的 M-奇异次微分.

注 8.4.5 注意,这里在定义 8.4.6 中使用了实值函数的上图象.在定义 8.4.2 中的集值映射退化为函数时,这两个定义是不一致的.当 f 在点 $\boldsymbol{x}^0$ 处是严格可微的,则有 $\partial^M f(\boldsymbol{x}^0) = \{\nabla f(\boldsymbol{x}^0)\}$.

注 8.4.6 设 f 在点 x 处是 Lipschitz 的,$\partial^0 f(\boldsymbol{x})$是 f 在 $\boldsymbol{x}$ 处的 G-次微分,则有

$$\partial^0 f(\boldsymbol{x}) = \mathrm{cl}\big(\mathrm{co}\big[\partial^M f(\boldsymbol{x}) + \partial^{\infty} f(\boldsymbol{x})\big]\big)$$

注 8.4.7 设凸集 $S \subset R^n$,对于 S 上的指示函数:

$$\delta_S(\boldsymbol{x}) = \begin{cases} 0 & \boldsymbol{x} \in S; \\ +\infty, & \boldsymbol{x} \in R^n \backslash S, \end{cases}$$

则有

$$\partial^M \delta_S(\boldsymbol{x}) = \partial^{\infty} \delta_S(\boldsymbol{x}) = \partial_S(\boldsymbol{x}) = N^M_S(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in S.$$

定理 8.4.4 设 $\boldsymbol{x} \in R^n$. 若向量函数 $\boldsymbol{f}: R^n \to R^m$ 在点 $\boldsymbol{x}$ 处是 Lipschitz 的,则

$$\partial^M \boldsymbol{f}(\boldsymbol{x})(\boldsymbol{y}^*) = \partial^M \langle \boldsymbol{y}^*,\ \boldsymbol{f} \rangle(\boldsymbol{x}) \quad \forall \boldsymbol{y}^* \in R^m.$$

证明 由定义 8.4.2 可得证.▯

为了讨论由定义 8.4.2 引进的 M-次微分的运算性质,以下给出集合的局部极点的概念.

定义 8.4.7 设 S_1、S_2 是闭集,点 $\boldsymbol{x}^0 \in S_1 \cap S_2$. 若存在点 $\boldsymbol{x}^0$ 的邻域 $U(\boldsymbol{x}^0)$及序列$\{\boldsymbol{a}^1_k\}$和$\{\boldsymbol{a}^2_k\}$, $\boldsymbol{a}^1_k \to 0$, $\boldsymbol{a}^2_k \to 0 (k \to \infty)$,使得

$$(S_1 - \boldsymbol{a}^1_k) \cap (S_2 - \boldsymbol{a}^2_k) \cap U(\boldsymbol{x}^0) = \varnothing \quad k = 1, 2, \cdots,$$

则称 $\boldsymbol{x}^0$ 是集系 $\{S_1, S_2\}$ 的**局部极点**. 若 $\{S_1, S_2\}$ 至少有一个局部极点，则称集合 S_1 和 S_2 形成一个**局部极点系**.

例如，设 $\boldsymbol{x}^0$ 是 S 的边界点，则 $\boldsymbol{x}^0$ 是 $\{\boldsymbol{x}^0, S\}$ 的局部极点. 又如，设函数 $f: R^n \to R$ 在点 $\boldsymbol{x}^0$ 处下半连续，$S \subset R^n$ 是闭凸集，$\boldsymbol{x}^0$ 是极小化问题

$$\min_{\boldsymbol{x} \in S} f(\boldsymbol{x})$$

的局部最优解. 取 $S_1 = \mathrm{epi} f$ 和 $S_2 = S \times \{f(\boldsymbol{x})\}$，则点 $(\boldsymbol{x}^0, f(\boldsymbol{x}^0))$ 是 $\{S_1, S_2\}$ 的局部极点.

依据上面所述，局部极点可以看成是局部最优解的推广.

下面给出局部极点的必要条件.

定理 8.4.5　设 S_1, S_2 是闭集. 若 $\bar{\boldsymbol{x}} \in S_1 \cap S_2$ 是 $\{S_1, S_2\}$ 的局部极点，则存在 $\boldsymbol{x}^* \in R^n$，使得

$$\boldsymbol{0} \neq \boldsymbol{x}^* \in N_{S_1}(\bar{\boldsymbol{x}}) \cap \left(-N_{S_2}(\bar{\boldsymbol{x}})\right). \tag{8.4.6}$$

证明　考虑序列 $\{\boldsymbol{a}_k^1\}$ 和 $\{\boldsymbol{a}_k^2\}$，$\boldsymbol{a}_k^1 \to \boldsymbol{0}$，$\boldsymbol{a}_k^2 \to \boldsymbol{0}$ $(k \to \infty)$，设 $U(\bar{\boldsymbol{x}})$ 是 $\bar{\boldsymbol{x}}$ 的邻域，则有

$$(S_1 - \boldsymbol{a}_k^1) \cap (S_2 - \boldsymbol{a}_k^2) \cap U(\bar{\boldsymbol{x}}) = \varnothing, \quad k = 1, 2, \cdots.$$

用 $d_S(\boldsymbol{x} + \boldsymbol{a})$ 表示点 $\boldsymbol{x} + \boldsymbol{a}$ 到集合 S 的距离，设

$$\rho_k(\boldsymbol{x}) = \left[d_{S_1}^2(\boldsymbol{x} + \boldsymbol{a}_k^1) + d_{S_2}^2(\boldsymbol{x} + \boldsymbol{a}_k^2)\right]^{1/2} + \|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2,$$

考虑一系列极小化问题：

$$\min_{\boldsymbol{x} \in R^n} \rho_k(\boldsymbol{x}), \quad k = 1, 2, \cdots. \tag{8.4.7}$$

因为对任何 k，$\rho_k(\boldsymbol{x})$ 关于 $\boldsymbol{x}$ 是连续的，并且水平集 $\{\boldsymbol{x} \mid \rho_k(\boldsymbol{x}) \geqslant 0\}$ 有界，所以对于每一个 k，存在(8.4.7)的最优解 $\boldsymbol{x}^k$. 设

$$\alpha_k = \left[d_{S_1}^2(\boldsymbol{x}^k + \boldsymbol{a}_k^1) + d_{S_2}^2(\boldsymbol{x}^k + \boldsymbol{a}_k^2)\right]^{1/2},$$

则有

$$\rho_k(x^k) = \alpha_k + \| x^k - \tilde{x} \|^2 \leqslant \rho_k(\tilde{x})$$
$$\leqslant [\| a_k^1 \|^2 + \| a_k^2 \|^2]^{1/2}.$$

由于已知当 $k \to \infty$ 时，有 $a_k^1 \to 0$ 和 $a_k^2 \to 0$，故得

$$\alpha_k \to 0^+ \quad 和 \quad x^k \to \tilde{x},\ k \to \infty. \tag{8.4.8}$$

设 $\omega_k^1 \in d_{S_1}(x^k + a_k^1)$ 和 $\omega_k^2 \in d_{S_2}(x^k + a_k^2)$，考虑极小化问题

$$\min_{x \in R^n} f_k(x) = [\| x + a_k^1 - \omega_k^1 \|^2 + \| x + a_k^2 - \omega_k^2 \|^2]^{1/2}$$
$$+ \| x - \tilde{x} \|^2 \quad k = 1, 2, \cdots.$$

显然，上式有最优解 $x^k (k = 1, 2, \cdots)$. 由 $f_k(x)\ (k = 1, 2, \cdots)$ 在 x 处是连续可微的，则

$$\nabla f_k(x^k) = x_k^{1*} + x_k^{2*} + 2(x^k - \tilde{x}) = 0, \tag{8.4.9}$$

其中 $x_k^{i*} = (x^k + a_k^i - \omega_k^i)/\alpha_k (i = 1, 2)$，并且 $\| x_k^{1*} \|^2 + \| x_k^{2*} \|^2 = 1$. 因此，存在 x_1^* 和 x_2^*，使得

$$\| x_1^* \|^2 + \| x_2^* \|^2 = 1,\ x_k^{1*} \to x_1^* \ 和\ x_k^{2*} \to x_2^*,\ k \to \infty.$$

对(8.4.9)取极限，根据定义 8.4.1 即得(8.4.6). □

注 8.4.8 若 S_1 和 S_2 是凸集，则(8.4.6)等价于凸集分离定理，即存在 $x^* \neq 0$，使得

$$\langle x^*, x^1 \rangle \leqslant \langle x^*, x^2 \rangle \quad \forall x^1 \in S_1,\ x^2 \in S_2.$$

事实上，由凸集分离性可以推出局部极点性. 在两个集合具有凸性时，凸集分离性与局部极点性是等价的.

定理 8.4.6 设集合 $\psi_1(x), \psi_2(x) \subset R^m (x \in R^n)$，集值映射 $\psi_1: R^n \to 2^{R^m}$ 和 $\psi_2: R^n \to 2^{R^m}$ 是闭的，$y^0 \in \psi_1(x^0) + \psi_2(x^0)$. 若集值映射 $\Theta: R^{n+m} \to 2^{R^{2m}}$，

$$\Theta(x, y) = \{(y_1, y_2) \in R^{2m} | y_1 \in \psi_1(x),$$
$$y_2 \in \psi_2(x),\ y_1 + y_2 = y\} \tag{8.4.10}$$

在(x^0, y^0)处局部有界，并且下式成立：

$$\partial^M \psi_1(x^0, y_1)(0) \cap \left(-\partial^M \psi_2(x^0, y_2)(0)\right) = \{0\}$$

$$\forall (y_1, y_2) \in \Theta(x^0, y^0), \qquad (8.4.11)$$

则对任意的 $y^* \in R^m$ 有

$$\partial^M(\psi_1 + \psi_2)(x^0, y^0)(y^*)$$

$$\subset \bigcup_{(y_1, y_2) \in \Theta(x, y)} \left[\partial^M \psi_1(x^0, y_1)(y^*) + \partial^M \psi_2(x^0, y_2)(y^*)\right].$$

$$(8.4.12)$$

证明　设 $x^* \in \partial^M(\psi_1 + \psi_2)(x^0, y^0)(y^*)$，由 ψ_1 和 ψ_2 是闭的可知，$\psi_1 + \psi_2$ 的图象也是闭的. 再由定义 8.4.2 和 F-法锥的关系得知，对任何一个正数列 $\{\varepsilon_k\}$，$\varepsilon_k \to 0^+ \ (k \to \infty)$，都可以找到序列 $\{x^k\}$ 和 $\{y^k\}$，$\{x_k^*\}$ 和 $\{y_k^*\}$，$x^k \to x^0$，$y^k \to y^0$，$x_k^* \to x^*$，$y_k^* \to y^* \ (k \to \infty)$，以及正数列 $\{\eta_k\}$，$\eta_k \to 0^+ \ (k \to \infty)$，使得 $y^k \in \psi_1(x^k) + \psi_2(x^k)$ 和对任意的$(x, y) \in \mathrm{graph}(\psi_1 + \psi_2)$，当 $\|x - x^k\| \leqslant \eta_k$，$\|y - y^k\| \leqslant 2\eta_k$ 时有

$$\langle x^k, x - x^k\rangle - \langle y_k^*, y - y^k\rangle$$

$$\leqslant \varepsilon_k\left(\|x - x^k\| + \|y - y^k\|\right). \qquad (8.4.13)$$

由于 $\Theta(x, y)$在(x^0, y^0)处是局部有界的，故存在序列$\{(y_k^1, y_k^2)\}$ $\subset \Theta(x^k, y^k)$收敛于 $(y_1^0, y_2^0) \in \Theta(x^0, y^0)$. 对每一个 $k = 1, 2, \cdots$，设

$$g_k^1(x, y_1, y_2) = \langle y_k^*, y_1 - y_k^1\rangle + \varepsilon_k\|y_1 - y_k^1\|$$

$$+ \delta\left((x_1, y_1), \mathrm{graph}\psi_1\right), \qquad (8.4.14)$$

$$g_k^2(x, y_1, y_2) = \langle x_k^*, x - x^k\rangle - \varepsilon_k\|x - x^k\|$$

$$- \langle y_k^*, y_2 - y_k^2\rangle - \varepsilon_k\|y_2 - y_k^2\|$$

$$- \delta\left((x, y_2), \mathrm{graph}\psi_2\right), \qquad (8.4.15)$$

由(8.4.13)和(8.4.14)作相应于 $R^n \times R^m \times R^m \times R$ 的闭子集：

$$S_k^1 = \mathrm{epi}g_k^1,\ S_k^2 = \mathrm{epi}g_k^2, \tag{8.4.16}$$

显然有

$$(\boldsymbol{x}^k,\ \boldsymbol{y}_k^1,\ \boldsymbol{y}_k^2,\ \boldsymbol{0}) \in S_k^1 \cap S_k^2, \quad k = 1, 2, \cdots.$$

由(8.4.13)得

$$S_k^1 \cap \left[S_k^2 - (\boldsymbol{0},\ \boldsymbol{0},\ \boldsymbol{0},\ v)\right] \cap \left[B((\boldsymbol{x}^k,\ \boldsymbol{y}_k^1,\ \boldsymbol{y}_k^2),\ \eta_k) \times R\right] = \varnothing$$

$$\forall v > 0,$$

故由定义 8.4.4 可知 $(\boldsymbol{x}^k,\ \boldsymbol{y}_k^1,\ \boldsymbol{y}_k^2,\ 0)$ 是$(S_k^1,\ S_k^2)$的局部极点. 由此，根据定理 8.4.5 得知，存在序列 $\{\boldsymbol{z}_k^*,\ \boldsymbol{y}_k^{1*},\ \boldsymbol{y}_k^{2*},\ \lambda_k\} \in R^{n+2m+1}$，使得

$$\left\|(\boldsymbol{z}_k^*,\ \boldsymbol{y}_k^{1*},\ \boldsymbol{y}_k^{2*},\ \lambda_k)\right\| = 1,\ k = 1, 2, \cdots \tag{8.4.17}$$

和

$$(\boldsymbol{z}_k^*,\ \boldsymbol{y}_k^{1*},\ \boldsymbol{y}_k^{2*},\ \lambda_k) \in N_{S_k^1}^M(\boldsymbol{x}^k,\ \boldsymbol{y}_k^1,\ \boldsymbol{y}_k^2,\ 0)$$

$$\cap \left[-N_{S_k^2}^M(\boldsymbol{x}_k,\ \boldsymbol{y}_k^1,\ \boldsymbol{y}_k^2,\ 0)\right].$$

由(8.4.14)和(8.4.16)得

$$(\boldsymbol{z}_k^*,\ \boldsymbol{y}_k^{1*},\ \boldsymbol{y}_k^{2*},\ \lambda_k) \in N_{\mathrm{epi}g_k^1}^M(\boldsymbol{x}^k,\ \boldsymbol{y}_k^1,\ \boldsymbol{y}_k^2,\ g_k^1(\boldsymbol{x}^k,\ \boldsymbol{y}_k^1,\ \boldsymbol{y}_k^2)). \tag{8.4.18}$$

根据定义 8.4.1 和定义 8.4.3，上式等价于

$$(\boldsymbol{z}_k^*,\ \boldsymbol{y}_k^{1*},\ \boldsymbol{y}_k^{2*}) \in \begin{cases} |\lambda_k|\partial^M g_k^1(\boldsymbol{x}^k,\ \boldsymbol{y}_k^1,\ \boldsymbol{y}_k^2), & \lambda_k \neq 0; \\ \partial^\infty g_k^1(\boldsymbol{x}^k,\ \boldsymbol{y}_k^1,\ \boldsymbol{y}_k^2), & \lambda_k = 0. \end{cases}$$

由此，从上式可以推得 $\boldsymbol{y}_k^{2*} = \boldsymbol{0}$，

$(\boldsymbol{z}_k^*,\ \boldsymbol{y}_k^{1*})$

$$\in \begin{cases} \left(0,\ |\lambda_k|(\boldsymbol{y}_k^* + \epsilon_k B)\right) + N_{\mathrm{graph}\phi_1}^M(\boldsymbol{x}^k,\ \boldsymbol{y}_k^1), & \lambda_k \neq 0; \\ N_{\mathrm{graph}\phi_1}^M(\boldsymbol{x}^k,\ \boldsymbol{y}_k^1), & \lambda_k = 0. \end{cases} \tag{8.4.19}$$

同样地，我们有

$$(-z^*, -y_k^{1*}, -y_k^{2*}, -\lambda_k) \in N^M_{\mathrm{epi}g_k^2}\big((x^k, y_k^1, y_k^2), g_k^2(x^k, y_k^1, y_k^2)\big).$$

从上式可以推得 $y_k^{1*} = \boldsymbol{0}$ 和

$$(z_k^*, y_k^{2*}) = \begin{cases} \Big(|\lambda_k|(x_k^* - \varepsilon_k B, -y_k^* - \varepsilon_k B)\Big) - N^M_{\mathrm{graph}\psi_2}(x^k, y_k^2)\Big), & \lambda_k \neq 0; \\ -N^M_{\mathrm{graph}\psi_2}(x^k, y_k^2), & \lambda_k = 0. \end{cases} \tag{8.4.20}$$

由(8.4.17)，$y_k^{1*} = \boldsymbol{0}$ 和 $y_k^{2*} = \boldsymbol{0}$，可选出序列 $\{z_k^*, \lambda_k\}$，使得当 $k \to \infty$ 时，它收敛于 $(z^*, \lambda) \in R^n \times R$：

$$\|(z^*, \lambda)\| = 1, \lambda \leqslant 0. \tag{8.4.21}$$

我们证明 $\lambda \neq 0$. 事实上，否则，若 $\lambda = 0$，则有 $z^* \neq \boldsymbol{0}$，由(8.4.21)，(8.4.19)和(8.4.20)，根据定义 8.4.2 我们有

$$z^* \in \partial^M\psi_1(x^0, y_1)(\boldsymbol{0}) \cap \big(-\partial^M\psi_2(x^0, y_2)(\boldsymbol{0})\big)$$

$$\forall (y_1^0, y_2^0) \in \Theta(x^0, y^0),$$

这导致与(8.4.11)矛盾. 因此，由(8.4.21)有 $\lambda < 0$. 对(8.4.19)取极限，得

$$z^*|\lambda|^{-1} \in \partial^M\psi_1(x^0, y_1^0)(y^*),$$

对(8.4.20)取极限，则得

$$(x^*, -z^*)|\lambda|^{-1} \in \partial^M\psi_2(x^0, y_2^0)(y^*).$$

由此证明了(8.4.12). ▯

推论 8.4.7 若对每一个 $(y_1, y_2) \in \Theta(x^0, y^0)$，集值映射 ψ_1 在 (x^0, y_1) 处或 ψ_2 在 (x^0, y_2) 处是伪 Lipschitz 的，Θ 在 (x^0, y^0) 处是局部有界的，则(8.4.12)成立.

证明 由定理 8.4.3 和定理 8.4.6 可以推得结论. ▯

推论 8.4.8 设向量函数 $\boldsymbol{f}_1: R^n \to R^m$ 在 $\boldsymbol{x}^0 \in R^n$ 处是可微的，集值映射 $\psi_2: R^n \to 2^{R^m}$ 是闭的，对任意的 $\boldsymbol{y}^0 \in \boldsymbol{f}_1(\boldsymbol{x}^0) + \psi_2(\boldsymbol{x}^0)$ 和 $\boldsymbol{y}^* \in R^m$，则有

$$\partial^M(\boldsymbol{f}_1 + \psi_2)(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{y}^*)$$
$$= [J\boldsymbol{f}_1(\boldsymbol{x}^0)]^* \boldsymbol{y}^* + \partial^M \psi_2(\boldsymbol{x}^0, \boldsymbol{y}^0 - \boldsymbol{f}_1(\boldsymbol{x}^0))(\boldsymbol{y}^*).$$

并且，$\boldsymbol{f}_1 + \psi_2$ 在 $(\boldsymbol{x}^0, \boldsymbol{y}^0)$ 处是正则和可微的当且仅当 ψ_2 在 $(\boldsymbol{x}^0, \boldsymbol{y}^0 - \boldsymbol{f}_1(\boldsymbol{x}^0))$ 处是正则和可微的.

证明 由定理 8.4.1 和定理 8.4.6 可得证. ▯

推论 8.4.9 设集值映射 $\psi_1: R^n \to 2^{R^m}$ 是闭的，又 $S \in R^n$ 是闭集，

$$\psi_2(\boldsymbol{x}) = \begin{cases} 0 \in R^m, & \boldsymbol{x} \in S; \\ \varnothing, & \boldsymbol{x} \notin S. \end{cases} \tag{8.4.22}$$

给定 $\boldsymbol{x}^0 \in S$，$\boldsymbol{y}^0 \in \psi_1(\boldsymbol{x}^0)$，设

$$\partial^M \psi_1(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{0}) \cap (-N_S^M(\boldsymbol{x}^0)) = \{\boldsymbol{0}\}, \tag{8.4.23}$$

则有

$$\partial^M(\psi_1 + \psi_2)(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{y}^*) \subset \partial^M \psi_1(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{y}^*) + N_S^M(\boldsymbol{x}^0)$$
$$\forall \boldsymbol{y}^* \in R^m. \tag{8.4.24}$$

再若 ψ_1 在 $\boldsymbol{x}^0$ 处是可微的或在 $(\boldsymbol{x}^0, \boldsymbol{y}^0)$ 处是正则和可微的，并且 S 在 $\boldsymbol{x}^0$ 处是正则的，则(8.4.24)成为等式.

证明 设 $\Theta(\boldsymbol{x}^0, \boldsymbol{y}^0) = \{(\boldsymbol{y}^0, \boldsymbol{0})\}$. 由(8.4.22)和(8.4.23)，根据定理 8.4.6 可推知(8.4.24)成立.

若 ψ_1 在 x^0 处是严格可微的，据推论 8.4.8 可知(8.4.24)为等式. 若 ψ_1 在 $(\boldsymbol{x}^0, \boldsymbol{y}^0)$ 处是正则和可微的，以及 S 在 $\boldsymbol{x}^0$ 处是正则的，则有

$$\partial^M \psi_1(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{y}^*) + N_S^M(\boldsymbol{x}^0)$$

$$= \partial^{MF}\psi_1(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{y}^*) + N_S^F(\boldsymbol{x}^0)$$

$$\subset \partial^{MF}(\psi_1 + \psi_2)(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{y}^*)$$

$$\subset \partial^{M}(\psi_1 + \psi_2)(\boldsymbol{x}^0, \boldsymbol{y}^0)(\boldsymbol{y}^*)$$

$$\forall \boldsymbol{y}^* \in R^m.$$

因此,(8.4.24)成为等式.▯

推论 8.4.10　设函数 $f_i: R^n \to R$ $(i = 1, 2)$ 是下半连续的. 若对 $\boldsymbol{x} \in R^m$,有

$$\partial^{\infty} f_1(\boldsymbol{x}) \cap \left(-\partial^{\infty} f_2(\boldsymbol{x})\right) = \{\boldsymbol{0}\}, \qquad (8.4.25)$$

则

$$\partial^{M}(f_1 + f_2)(\boldsymbol{x}) \subset \partial^{M} f_1(\boldsymbol{x}) + \partial^{M} f_2(\boldsymbol{x}), \quad (8.4.26)$$

$$\partial^{\infty}(f_1 + f_2)(\boldsymbol{x}) \subset \partial^{\infty} f_1(\boldsymbol{x}) + \partial^{\infty} f_2(\boldsymbol{x}). \quad (8.4.27)$$

再若 f_1 或 f_2 在 $\boldsymbol{x}$ 处是可微的,则(8.4.26)和(8.4.27)成为等式.

证明　由定理 8.4.4 和定理 8.4.6 可得证.▯

推论 8.4.11　设 S_1, S_2 是闭集. 若 $\boldsymbol{x} \in S_1 \cap S_2$, $N_{S_1}^M(\boldsymbol{x}) \cap \left(-N_{S_2}^M(\boldsymbol{x})\right) = \{\boldsymbol{0}\}$, 则

$$N_{S_1 \cap S_2}^M(\boldsymbol{x}) \subset N_{S_1}^M(\boldsymbol{x}) + N_{S_2}^M(\boldsymbol{x}).$$

证明　由定义 8.4.1,例 8.4.1 和定理 8.4.6 可以推得.▯

下面讨论复合集值映射的 M-次微分问题.

设集合 $\theta(\boldsymbol{x}) \subset R^m (\boldsymbol{x} \in R^n)$, $\psi(\boldsymbol{y}) \subset R^q (\boldsymbol{y} \in R^m, q > 1)$, 对于集值映射 $\theta: R^n \to 2^{R^m}$ 和 $\psi: R^m \to 2^{R^q}$, 考虑复合集值映射

$$(\psi \circ \theta): R^n \to 2^{R^q},$$

$$\boldsymbol{x} \mapsto (\psi \circ \theta)(\boldsymbol{x}) = \psi\left(\theta(\boldsymbol{x})\right) = \bigcup_{\boldsymbol{y} \in \theta(\boldsymbol{x})} \psi(\boldsymbol{y}).$$

定理 8.4.12　设 $\boldsymbol{x}^0 \in R^n$, graphθ 和 graphψ 是闭的, $\boldsymbol{z}^0 \in (\psi \circ \theta)(\boldsymbol{x}^0)$. 考虑集值映射 $\Xi: R^n \times R^q \to 2^{R^m}$, $(\boldsymbol{x}, \boldsymbol{z}) \mapsto \Xi(\boldsymbol{x}, \boldsymbol{z})$,

$$\Xi(\boldsymbol{x}, \boldsymbol{z}) = \theta(\boldsymbol{x}) \cap \psi^{-1}(\boldsymbol{z}) = \{\boldsymbol{y} \in \theta(\boldsymbol{x}) \mid \boldsymbol{z} \in \theta(\boldsymbol{y})\}. \tag{8.4.28}$$

若 Ξ 在点$(\boldsymbol{x}^0, \boldsymbol{z}^0)$处是局部有界的,且条件

$$\partial^M\psi(\boldsymbol{y}, \boldsymbol{z}^0)(\boldsymbol{0}) \cap \mathrm{Ker}\partial^M\theta(\boldsymbol{x}^0, \boldsymbol{y}) = \{\boldsymbol{0}\}$$

$$\forall \boldsymbol{y} \in \theta(\boldsymbol{x}^0) \cap \psi^{-1}(\boldsymbol{z}^0) \tag{8.4.29}$$

成立,则

$$\partial^M(\psi \circ \theta)(\boldsymbol{x}^0, \boldsymbol{z}^0)$$

$$\subset \bigcap_{\boldsymbol{y} \in \theta(\boldsymbol{x}^0) \cap \psi^{-1}(\boldsymbol{x}^0)} \left[\partial^M\theta(\boldsymbol{x}^0, \boldsymbol{y}) \circ \partial^M\psi(\boldsymbol{y}, \boldsymbol{z}^0)\right]. \tag{8.4.30}$$

证明 设 $\boldsymbol{x}^* \in \partial^M(\psi \circ \theta)(\boldsymbol{x}^0, \boldsymbol{z}^0)(\boldsymbol{z}^*)$,我们证明存在 $\boldsymbol{y} \in \theta(\boldsymbol{x}^0) \cap \psi^{-1}(\boldsymbol{z}^0)$ 和 $\boldsymbol{y}^* \in \partial^M\psi(\boldsymbol{y}, \boldsymbol{z}^0)(\boldsymbol{z}^*)$,使得 $\boldsymbol{x}^* \in \partial^M\theta(\boldsymbol{x}^0, \boldsymbol{y})(\boldsymbol{y}^*)$. 为此,作集值映射 $g: R^n \times R^m \to 2^{R^q}$,$(\boldsymbol{x}, \boldsymbol{y}) \mapsto g(\boldsymbol{x}, \boldsymbol{y})$,

$$g(\boldsymbol{x}, \boldsymbol{y}) = \psi(\boldsymbol{y}) + \Delta^q((\boldsymbol{x}, \boldsymbol{y}), \mathrm{graph}\theta), \tag{8.4.31}$$

其中

$$\Delta^q((\boldsymbol{x}, \boldsymbol{y}), \mathrm{graph}\theta) = \begin{cases} 0, & (\boldsymbol{x}, \boldsymbol{y}) \in \mathrm{graph}\theta; \\ \varnothing, & (\boldsymbol{x}, \boldsymbol{y}) \notin \mathrm{graph}\theta. \end{cases}$$

对任意的 $\boldsymbol{z}^* \in R^q$ 我们证明

$$\partial^M(\psi \circ \theta)(\boldsymbol{x}^0, \boldsymbol{z}^0)(\boldsymbol{z}^*)$$

$$\subset \left\{\boldsymbol{x}^* \,\middle|\, (\boldsymbol{x}^*, \boldsymbol{0}) \in \bigcup_{\boldsymbol{y} \in \theta(\boldsymbol{x}^0) \cap \psi^{-1}(\boldsymbol{x}^0)} \partial^M g(\boldsymbol{x}^0, \boldsymbol{y}, \boldsymbol{z}^0)(\boldsymbol{z}^*)\right\}. \tag{8.4.32}$$

设 $\boldsymbol{x}^* \in \partial^M(\psi \circ \theta)(\boldsymbol{x}^0, \boldsymbol{z}^0)(\boldsymbol{z}^*)$,由于 Ξ 在点$(\boldsymbol{x}^0, \boldsymbol{z}^0)$处是局部有界的,则 graph$(\psi \circ \theta)$是闭的. 根据定义 8.4.2 和(8.4.3)式可知,存在序列$\{\boldsymbol{x}^k\}$和$\{\boldsymbol{z}^k\}$,以及$\{\boldsymbol{x}_k^*\}$和$\{\boldsymbol{z}_k^*\}$,$\boldsymbol{x}^k \to \boldsymbol{x}^0$,$\boldsymbol{z}^k \to \boldsymbol{z}^0$,以及 $\boldsymbol{x}_k^* \to \boldsymbol{x}^*$,$\boldsymbol{z}_k^* \to \boldsymbol{z}^*$ $(k \to \infty)$,使得

$$\boldsymbol{z}^k \in (\psi \circ \theta)(\boldsymbol{x}^k) \text{ 和 } (\boldsymbol{x}_k^*, -\boldsymbol{z}_k^*) \in N^M_{\mathrm{graph}(\psi \circ \theta)}(\boldsymbol{x}^k, \boldsymbol{z}^k),$$

$$k = 1, 2, \cdots.$$

取任一序列$\{y^k\}$, $y^k \in \theta(x^k) \cup \psi^{-1}(z^k)$，又由于 Ξ 在点(x^0, z^0)处是局部有界的，故不妨设 $y_k^0 \to y^0 \in \Xi(x^0, z^0)$. 由(8.4.2)令 $\epsilon = 0$，我们有

$$\gamma = \limsup_{\substack{(x,y,z)\to(x^k,y^k,z^k)\\ y\in\theta(x),\, z\in\psi(y)}} \frac{\langle x_k^*, x - x^k\rangle - \langle z_k^*, z - z^k\rangle}{\|(x, y, z) - (x^k, y^k, z^k)\|}$$

$$\leqslant \limsup_{\substack{(x,z)\to(x^k,z^k)\\ z\in(\psi\cdot\theta)(x)}} \frac{\langle x_k^*, x - x^k\rangle - \langle z_k^*, z - z^k\rangle}{\|(x, z) - (x^k, z^k)\|} \leqslant 0.$$

由此，得知 $(x_k^*, \boldsymbol{0} - z_k^*) \in N^M_{\mathrm{graph}_R}(x^k, y^k, z^k)$，再由(8.4.3)得到(8.4.32). 现在由(8.4.31)利用推论 8.4.9，即可推得(8.4.30)成立. □

推论 8.4.13　在定理 8.4.12 中，条件(8.4.31)可以换成：对任意的 $y \in \theta(x^0) \cup \psi^{-1}(z^0)$，$\Xi$ 在(y, z^0)处是伪 Lipschitz 的.

证明　由定理 8.4.3 和定理 8.4.12 可得证. □

推论 8.4.14　设向量函数 $f: R^m \to R^q$ 在点 $y \in \theta(x^0) \cap f^{-1}(z^0)$ 处是可微的，graphθ 是闭的. 若由(8.4.28)定义的 Ξ 在点(x^0, z^0)处是局部有界的，则

$$\partial^M(f \circ \theta)(x^0, z)(z^*)$$

$$\subset \bigcup_{y\in\theta(x^0)\cap f^{-1}(z^0)} \left[\partial^M\theta(x^0, y) \circ (Jf(y))^*\right].$$

证明　由定理 8.4.1 和定理 8.4.12 可以得证. □

推论 8.4.15　设向量函数 $f: R^m \to R^m$ 在点 x^0 处是可微的，graphψ 是闭的. 若

$$\partial^M\psi\big(f(x^0), z^0\big)(\boldsymbol{0}) \cap \mathrm{Ker}(Jf(x^0))^* = \{\boldsymbol{0}\},$$

则对任意的 $z^* \in R^q$ 有

$$\partial^M(\psi\circ f)(x^0, z^0)(z^*)$$

$$\subset \big(Jf(x^0)\big)^* \partial^M\psi(f(x^0), z^0)(z^*).$$

证明 由定理 8.4.1 和定理 8.4.12 可得证.▯

定理 8.4.16 设向量函数 $\theta: R^n \to R^m$ 在 x^0 处是 Lipschitz 连续的,集值映射 $\psi: R^m \to 2^{R^q}$ 是闭的,则由(8.4.31)定义的 g 使(8.4.32)成为等式.再若条件(8.4.29)成立,$y^0=\theta(x^0)$,则

$$\partial^M(\psi\circ\theta)(x^0, z^0)(z^*)$$

$$\subset \bigcup_{y^*\in\partial^M\psi(\theta(x^0), z^0)(z^*)} \partial^M\psi\langle y^*, \theta\rangle(x^0). \tag{8.4.33}$$

(1) 若 ψ 在 $y^0=\theta(x^0)$ 处是可微的,则

$$\partial^M(\psi\circ\theta)(x^0, z^0)(z^*)$$

$$=\partial^M\big\langle (J\psi(y^0))^* z^*, \theta\big\rangle(x^0) \quad \forall z^*\in R^q. \tag{8.4.34}$$

(2) 若 θ 在 x^0 处是可微的,ψ 在$(\theta(x^0), z^0)$处是可微和正则的,则有

$$\partial^M(\psi\circ\theta)(x^0, z^0)(z^*)$$

$$=\big(J\theta(x^0)\big)^*\partial^M\psi(\theta(x^0), z^0)(z^*) \quad \forall z^*\in R^q, \tag{8.4.35}$$

并且 $\theta\circ\psi$ 在(x^0, z^0)处是可微和正则的.

证明 先证明(8.4.32)的反包含情形:设$(x^*, \boldsymbol{0})$属于(8.4.32)的右端,即

$$(x^*, \boldsymbol{0}, -z^*)\in N^M_{\text{graph}g}(x^0, \theta(x^0), z^0).$$

由(8.4.3)可知,存在序列$\{x^k\}$、$\{z^k\}$、$\{x_k^*\}$、$\{y_k^*\}$和$\{z_k^*\}$,$x^k\to x^0$,$z^k\to z^0$,$x_k^*\to x^*$,$y_k^*\to\boldsymbol{0}$和 $z_k^*\to z^*(k\to\infty)$,使得

$$\limsup_{\substack{(x,z)\to(x^k,z^k)\\ z\in(\psi\circ\theta)(x)}} \frac{\langle (x_k^*, y_k^*, -z_k^*), (x, \theta(x), z)-(x^k, \theta(x^k), z^k)\rangle}{\|(x, \theta(x), z)-(x^k, \theta(x^k), z^k)\|}\leqslant 0,$$

$$k=1, 2, \cdots. \tag{8.4.36}$$

设 γ 为 θ 在点 x^0 处的 Lipschitz 常数，则有

$$\|(x, \theta(x), z) - (x^k, \theta(x^k), z^k)\|$$
$$\leqslant (\gamma + 1)(\|x - x^k\| + \|z - z^k\|).$$

因此有

$$\frac{\langle (x_k^*, -z_k^*), (x, z) - (x^k, z^k)\rangle}{\|(x, z) - (x^k, z^k)\|}$$

$$\leqslant (\gamma + 1)\frac{\langle (x_k^*, -z_k^*), (x, z) - (x^k, z^k)\rangle}{\|(x, \theta(x), z) - (x^k, \theta(x^k), z^k)\|}$$

$$\leqslant (\gamma + 1)\frac{\langle (x_k^*, y_k^*, -z_k^*), (x, \theta(x), z) - (x^k, \theta(x^k), z^k)\rangle}{\|(x, \theta(x), z) - (x^k, \theta(x^k), z^k)\|}$$

$$+ (\gamma + 1)\|y_k^*\|.$$

设 $\varepsilon_k = (\gamma + 1)\|y_k^*\|$，则(8.4.36)为

$$\limsup_{\substack{(x, z)\to(x^k, z^k)\\ z\in(\psi\circ\theta)(x)}} \frac{\langle x_k^*, x - x^k\rangle - \langle z_k^*, z - z^k\rangle}{\|(x, z) - (x^k, z^k)\|} \leqslant \varepsilon_k,\ k = 1, 2, \cdots.$$

上式等价为 $(x_k^*, -z_k^*) \in N^F_{\mathrm{graph}(\psi\circ\theta)}((x_k, z_k), \varepsilon_k)$. 令 $\varepsilon_k \to 0^+$，由(8.4.3)得 $x^* \in \partial^M(\psi\circ\theta)(x^0, z^0)(z^*)$，再由定理 8.4.4 得到(8.4.33).

(1) 由推论 8.4.7 可得(8.4.34)成立.

(2) 由推论 8.4.8 可推得(8.4.35)成立. ▯

推论 8.4.17　设 θ 和 ψ 都是单值映射(向量函数). 若它们分别在 x^0 和 $\theta(x^0)$ 处是 Lipschitz 连续的，则对任意的 $z^* \in R^q$ 有

$$\partial^M\langle z^*, \psi\circ\theta\rangle(x^0) \subset \bigcup_{y^*\in\partial^M\langle z^*, \psi\rangle(\theta(x^0))} \partial^M\langle y^*, \theta\rangle(x^0).$$

再若 ψ 在 $\theta(x^0)$ 处是严格可微的，则上式成为等式.

证明　由定理 8.4.16 可推证. ▯

推论 8.4.18 设θ和ψ与推论 8.4.17 中的相同，则对任意的$z^* \in R^q$有

$$\left(J(\psi \circ \theta)(x^0)\right)^* z^* \subset \mathrm{co}\left\{ \bigcup_{y^* \in \partial^M(z^*, \psi)(\theta(x^0))} \partial^M(y^*, \theta(x^0)) \right\}$$

$$\subset \mathrm{co}\left\{ \bigcup_{y^* \in (J\psi(\theta(x^0)))^* z^*} \partial^M(y^*, \theta)(x^0) \right\},$$

$$\left\{J(\psi \circ \theta)(x^0)^*\right\} z^* \subset \mathrm{co}\left\{(J\theta(x^0))^* \partial^M(z^*, \psi)(\theta(x^0))\right\}$$

$$\subset \mathrm{co}\left\{(J\psi(\theta(x^0))J\theta(x^0))^* z^*\right\},$$

若ψ在$\theta(x^0)$处是可微的，则

$$\left(J(\psi \circ \theta)(x^0)\right)^* z^* = \left(J\psi(\theta(x^0))J\theta(x^0)\right)^* z^*.$$

证明 由定理 8.4.16 不难推证. ▯

现在讨论乘法规则. 设集合$\theta_1(x)$、$\theta_2(x) \subset R^m (x \in R^n)$，$\theta_1: R^n \to 2^{R^m}$和$\theta_2: R^n \to 2^{R^m}$是集值映射，记

$$\langle \theta_1, \theta_2 \rangle (x) = \bigcup_{y_1 \in \theta_1(x),\, y_2 \in \theta_2(x)} \langle y_1, y_2 \rangle.$$

定理 8.4.19 设$x \in R^n$，$\mathrm{graph}\theta_1$和$\mathrm{graph}\theta_2$是闭的，$z \in \langle \theta_1, \theta_2 \rangle (x)$. 若集值映射$\Xi: R^n \times R \to 2^{R^m}$，$(x, z) \mapsto \Xi(x, z)$，

$$\Xi(x, z) = \left\{(y^1, y^2) \in R^{2m} \mid y^1 \in \theta_1(x),\right.$$

$$\left. y^2 \in \theta_2(x),\ \langle y^1, y^2 \rangle = z \right\}$$

在(x, z)处是局部有界的，则对任意的$\alpha \in R$，有

$$\partial^M \langle \theta_1, \theta_2 \rangle (x, z)(\alpha)$$

$$\subset \bigcup_{(y^1, y^2) \in \Xi(x, z)} \partial^M(\theta_1, \theta_2)(x, y^1, y^2)(\alpha y^2, \alpha y^1). \qquad (8.4.37)$$

若θ_1和θ_2是单值映射，并且在点x处是 Lipschitz 的，则

$$\partial^M \langle \theta_1, \theta_2 \rangle (x)$$

$$= \partial^M\big(\langle\theta_2(\boldsymbol{x}),\ \theta_1\rangle + \langle\theta_1(\boldsymbol{x}),\ \theta_2\rangle\big)(\boldsymbol{x}). \tag{8.4.38}$$

证明　将集值映射$\langle\theta_1,\ \theta_2\rangle$看作下面两个集值映射的复合：

$$\theta(\boldsymbol{x}) = \big(\theta_1(\boldsymbol{x}),\ \theta_2(\boldsymbol{x})\big):R^n \to 2^{R^{2m}},$$

$$\psi(\boldsymbol{y}^1,\ \boldsymbol{y}^2) = \langle\boldsymbol{y}^1,\ \boldsymbol{y}^2\rangle:R^{2m} \to 2^R,$$

即

$$\langle\theta_1,\ \theta_2\rangle(\boldsymbol{x}) = (\psi \circ \theta)(\boldsymbol{x}).$$

根据推论 8.4.14 得知(8.4.37)成立. 再由定理 8.4.16 和定理 8.4.4 推得(8.4.38)成立. ▯

推论 8.4.20　设 θ_1 或者 θ_2 是单值映射(向量函数),并且在点 $\boldsymbol{x} \in R^n$ 处是 Lipschitz 的,则

$$\partial^M\langle\theta_1,\ \theta_2\rangle(\boldsymbol{x}) \subset \partial^M\langle\theta_2(\boldsymbol{x}),\ \theta_1\rangle(\boldsymbol{x}) + \partial^M\langle\theta_1(\boldsymbol{x}),\ \theta_2\rangle(\boldsymbol{x})$$

$$\subset \big(J\theta_1(\boldsymbol{x})\big)^*\theta_2(\boldsymbol{x}) + \big(J\theta_2(\boldsymbol{x})\big)^*\theta_1(\boldsymbol{x}).$$

再若 θ_1 或 θ_2 在 $\boldsymbol{x}$ 处是可微的,则上式成为等式.

证明　由定理 8.4.1 和定理 8.4.19 可推得. ▯

定理 8.4.21　题设与定理 8.4.19 相同. 若

$$\partial^M\theta_1(\boldsymbol{x},\ \boldsymbol{y}^1)(\boldsymbol{0}) \cap \big(-\partial^M\theta_2(\boldsymbol{x},\ \boldsymbol{y}^2)(\boldsymbol{0})\big) = \{\boldsymbol{0}\}$$

$$\forall(\boldsymbol{y}^1,\ \boldsymbol{y}^2) \in \Xi(\boldsymbol{x},\ z),$$

则对任何实数 α 有

$$\partial^M\langle\theta_1,\ \theta_2\rangle(\boldsymbol{x},\ \boldsymbol{z})(\alpha)$$

$$\subset \bigcup_{(\boldsymbol{y}^1,\ \boldsymbol{y}^2)\in\Xi(\boldsymbol{x},\ z)}\big[\partial^M\theta_1(\boldsymbol{x},\ \boldsymbol{y}^1)(\alpha\boldsymbol{y}^2) + \partial^M\theta_2(\boldsymbol{x},\ \boldsymbol{y}^2)(\alpha\boldsymbol{y}^1)\big]. \tag{8.4.39}$$

证明　作复合映射:

$$\langle\theta_1,\ \theta_2\rangle(\boldsymbol{x}) = (\psi \circ \theta)(\boldsymbol{x}),$$

$$\theta(\boldsymbol{x}) = (\boldsymbol{x}, \boldsymbol{x}): R^n \to 2^{R^{2n}},$$

$$\psi(\boldsymbol{x}^1, \boldsymbol{x}^2) = \langle \theta_1(\boldsymbol{x}^1), \theta_2(\boldsymbol{x}^2) \rangle: R^{2n} \to 2^R.$$

利用推论 8.4.14 可以推得(8.4.39).▯

定理 8.4.22 设集合 $\theta_1(\boldsymbol{x})$、$\theta_2(\boldsymbol{x}) \subset R(\boldsymbol{x} \in R^n)$, $\theta_1, \theta_2: R^n \to 2^R$ 是集值映射, $\mathrm{graph}\theta_1$ 和 $\mathrm{graph}\theta_2$ 是闭凸的. 考虑集值映射:

$$\left(\frac{\theta_1}{\theta_2}\right)(\boldsymbol{x}) = \bigcup_{y_1 \in \theta_1(\boldsymbol{x}),\ y_2 \in \theta_2(\boldsymbol{x})} \{y_1/y_2 \mid y_2 \neq 0\}.$$

设 $z \in \left(\frac{\theta_1}{\theta_2}\right)(\boldsymbol{x})$. 若集值映射 $\Xi: R^n \times R \to 2^{R^2}$, $(\boldsymbol{x}, z) \mapsto \Xi(\boldsymbol{x}, z)$,

$$\Xi(\boldsymbol{x}, z) = \{(y_1, y_2) \in R^2 \mid y_1 \in \theta_1(\boldsymbol{x}),$$
$$y_2 \in \theta_2(\boldsymbol{x}), y_1/y_2 = z\}$$

在$(\boldsymbol{x}, z)$处是局部有界的,则对任意的 $\alpha \in R$ 有

$$\partial^M\left(\frac{\theta_1}{\theta_2}\right)(\boldsymbol{x}, z)(\alpha) \subset \bigcup_{(y_1, y_2) \in \Xi(\boldsymbol{x}, z)} \Big[\partial^M(\theta_1, \theta_2)$$
$$\times (\boldsymbol{x}, y_1, y_2)(\alpha y_1 - \alpha y_2)\Big]\Big/(y_2)^2. \qquad (8.4.40)$$

若 θ_1 和 θ_2 在 $\boldsymbol{x}$ 处是单值和 Lipschitz 的,则对任意的 $\alpha \in R$ 有

$$\partial^M\left(\frac{\theta_1}{\theta_2}\right)(\boldsymbol{x}, z)(\alpha) \subset \bigcup_{(y_1, y_2) \in \Xi(\boldsymbol{x}, z)} \Big[\partial^M\theta_1(\boldsymbol{x}, y_1)(\alpha y_2)$$
$$\times \partial^M\theta_2(\boldsymbol{x}, y_2)(-\alpha y_1)\Big]\Big/(y_2)^2. \qquad (8.4.41)$$

证明 作复合集值映射: $\left(\frac{\theta_1}{\theta_2}\right)(\boldsymbol{x}) = (\psi \circ \theta)(\boldsymbol{x})$, 其中 $\theta(\boldsymbol{x}) = (\theta_1(\boldsymbol{x}), \theta_2(\boldsymbol{x}))$ 和 $\theta(y_1, y_2) = y_1/y_2$. 由推论 8.4.14 可以得到(8.4.40). 若在复合集值映射 $\left(\frac{\theta_1}{\theta_2}\right)(\boldsymbol{x}) = (\psi \circ \theta)(\boldsymbol{x})$ 中, $\theta(\boldsymbol{x}) =$

$(\boldsymbol{x}, \boldsymbol{x})$, $\psi(y_1, y_2) = \theta_1(\boldsymbol{x})/\theta_2(\boldsymbol{x})$, 利用推论 8.4.15 即可以得到 (7.4.41). ▯

容易推得下面结论.

定理 8.4.23　设集合 $\theta_1(\boldsymbol{x}), \theta_2(\boldsymbol{x}) \subset R(\boldsymbol{x} \in R^n)$, 集值映射 $\theta_1: R^n \to 2^R$ 和 $\theta_2: R^n \to 2^R$ 是闭的. 考虑集值映射 $\theta_1 \vee \theta_2: R^n \to 2^R$, $\boldsymbol{x} \mapsto (\theta_1 \vee \theta_2)(\boldsymbol{x})$,

$$(\theta_1 \vee \theta_2)(\boldsymbol{x}) = \{z \in R \mid \exists\ y_1 \in \theta_1(\boldsymbol{x}),\ y_2 \in \theta_2(\boldsymbol{x}),$$
$$\max\{y_1, y_2\} = z\}.$$

设 $\boldsymbol{x} \in R^n$, $z \in (\theta_1 \vee \theta_2)(\boldsymbol{x})$. 若集值映射 $\Xi: R^n \times R \to 2^{R^2}$, $(\boldsymbol{x}, z) \mapsto \Xi(\boldsymbol{x}, z)$,

$$\Xi(\boldsymbol{x}, z) = \{(y_1, y_2) \in R^2 \mid y_1 \in \theta_1(\boldsymbol{x}),\ y_2 \in \theta_2(\boldsymbol{x}),$$
$$\max\{y_1, y_2\} = z\}$$

在 $(\boldsymbol{x}, z)$ 处是局部有界的, 则对任意的 $\alpha \in R$ 有

$$\partial^M(\theta_1 \vee \theta_2)(\boldsymbol{x}, z)(\alpha) \subset \bigcup_{\substack{(y_1, y_2) \in \Xi(\boldsymbol{x}, z) \\ \alpha \in \Gamma(\alpha)}} \big[\partial^M(\theta_1, \theta_2)$$
$$\times(\boldsymbol{x}, y_1, y_2)(\lambda\alpha, (1-\lambda)\alpha)\big],$$

其中 $\Gamma(\alpha) = [0, 1]$(当 $\alpha \geqslant 0$) 和 $\Gamma(\alpha) = \{0, 1\}$(当 $\alpha < 0$). 再若有

$$\partial^M\theta_1(\boldsymbol{x}, y_1)(0) \cap (-\partial^M\theta_2(\boldsymbol{x}, y_2)(\boldsymbol{0}) = \{\boldsymbol{0}\}$$
$$\forall (y_1, y_2) \in \Xi(\boldsymbol{x}, z),$$

则

$$\partial^M(\theta_1 \cup \theta_2)(\boldsymbol{x}, z)(\alpha)$$
$$\subset \bigcup_{\substack{(y_1, y_2) \in \Xi(\boldsymbol{x}, z) \\ \alpha \in \Gamma(\alpha)}} \big[\partial^M\theta_1(\boldsymbol{x}, y_1)(\lambda\alpha) + \partial^M\theta_2(\boldsymbol{x}, y_2)(1-\lambda)\alpha)\big].$$

参考文献

[1] Köthe G. Topologische Lineare Räume. Berlin-Gotting: Springer-Verlag, 1960

[2] Jameson G J O. Ordered Linear Space. Lecture Notes in Math, 104, Berlin: Springer-Verlag, 1970

[3] Rockafellar R T. Convex Analysis. Princeton Mathematics Ser, Vol 28, Princeton: Princeton Univ. Press, 1970

[4] Clarke F H. Necessary conditions for nonsmooth problems in optimal control and the calculus of variations. Ph. D thesis. Univ. of Washington, 1973

[5] Ekeland I et Teman R. Analyse Convexe et Problémes Variationnels. Dunod-Gauthier-Villars, 1974

[6] Bartle R G. The Elements of Real Analysis. 2nd ed. New York: John Wiley & Sons, 1976

[7] Пшеничный Б Н. Выпуклый Анализ и Экстремальные Задачи. Москва: Наука, 1980

[8] Tanino T, Sawaragi Y. Conjugate maps and duality in multiobjective optimization. J. of Opti. Theory and Appl., 1980, 31(4): 473～499

[9] Ioffe A D. Nonsmooth analysis: differential calculus of nondifferentiable mappings. Trans. Amer. Math. Soc., 1981, 266:1～56

[10] Канторович Л В. 泛函分析(上册). 刘证等译. 北京:高等教育出版社, 1982

[11] Clarke F H. Optimization and Nonsmooth Analysis. New York: John Wiley & Sons-Interscience, 1983

[12] Köthe G. Topological Vector Spaces I. Berlin: Springer-Verlag, 1983

[13] Kanniappan P. Necessary conditions for optimality of nondifferentiable convex multiobjective programming. J. of Opti. Theory and Appl., 1983, 40(2):167～174

[14] Minami M. Weak Pareto-optimal necessary conditions in a nondifferentiable multiobjective programming on a Banach space. J. of Opti. Theory and Appl., 1983, 41(3): 451～461

[15] Jan Van Tiel. Convex Analysis, An Introductory Text, John Wiley & Sons, 1984

[16] Aubin J P, Ekeland I. Applied Nonlinear Analysis. New York: John Wiley & Sons, 1984

[17] Aubin J P. Lanalyse non Lineaire et ses Motivations Economiques. Paris: Masson, 1984

[18] 夏道行,杨亚立.线性拓扑空间引论.上海:上海科学技术出版社,1986

[19] 张恭庆,林源渠.泛函分析讲义(上册).北京:北京大学出版社,1987

[20] 夏道行等.泛函分析第二教程.北京:高等教育出版社,1987

[21] Kerysznig E. 泛函分析导论及应用.蒋正新译.北京:北京航空学院出版社,1987

[22] Ferro F. A minimax theorem for vector-valued functions. J. of Opti. Theory and Appl., 1989, 60(1): 19~31

[23] Luc D T. Theory of Vector Optimization. Berlin: Springer Verlag, 1989

[24] 胡毓达,胡一凡.锥拟凸与拓扑向量空间多目标最优化有效解集和弱有效解集的连通性.应用数学学报,1989, 12(1):115~123

[25] 史树中.凸分析.上海:上海科学技术出版社,1990

[26] 胡毓达,徐永明.扰动多目标规划的次微分稳定性.数学学报,1992, 35(5):577~586

[27] 胡毓达.多目标规划有效性理论.上海:上海科学技术出版社,1994

[28] Morduckhovich B S. Generalized differential calculus for nonsmooth and set-valued mappings. J. of Math. Anal. and Appl., 1994, 183: 250~288

[29] Lai-Jiu Lin. Optimization of set-valued function. J. of Math. Anal. and Appl., 1994, 186(1):30~51

[30] Morduckhovich B S. Stability theory for parametric generalized equations and variational inequalities via nonsmooth analysis. Trans. Amer Math Soc. 1994, 343(2): 609~657

[31] Tanaka T. Generalized Quasiconvexities, Cone saddle points, and minimax theorem for vector-valued functions. J. of Opti. Theory and Appl., 1994, 81(2):355~377

[32] 周性伟.实变函数.北京:科学出版社,1998

[33] Hu Yuda, Meng Zhiqing. The existence theorem of the cone-weak subdifferential of set-valued mapping. Appl. Math. A J. of Chinese Univ., 1998, 13B(4):411～415

[34] 孟志青. 集值函数 Hahn-Banach 定理. 应用数学和力学, 1998, 19(1): 55～61

[35] 孟志青, 唐勇. 集值函数的一种广义微分的存在性. 应用数学, 1998, 11(4):99～101

[36] 胡毓达、孟志青. 拓扑向量空间集值映射的锥有效次微分及其存在性. 上海交通大学学报, 2000, 34(8):1114～1117

[33] Hu Yuda, Meng Zhiqing. The existence theorem of the cone-weak subdifferential of set-valued mapping. Appl. Math. A J. of Chinese Univ., 1998, 13B(4):411~415

[34] 孟志青. 集值函数 Hahn-Banach 定理. 应用数学和力学, 1998, 19(1): 55~61

[35] 孟志青, 唐勇. 集值函数的一种广义微分的存在性. 应用数学, 1998, 11(4): 99~101

[36] 胡毓达、孟志青. 拓扑向量空间集值映射的锥有效次微分及其存在性. 上海交通大学学报, 2000, 34(8): 1114~1117